

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 08:29:25, Search time 40.6 Seconds
(without alignments)
534,882 Million cell updates/sec

Title: US-09-598-062-6
Perfect score: 1245
Sequence: FENINWVEAIVNIENQF FENINWVEAIVNIENQF 229

Scoring table: BLASDM62
Gapop 10.0, Gapext 0.5

Searched: 28138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 28138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: p12_71-★
1: p12:★
2: p12:★
3: p13:★
4: p14:★

Pred. loc. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is primarily for analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	length	ID	Description
1	509	40.9	454	JC4848	cysteine proteinase
2	496	39.8	376	F87436	cysteine proteinase
3	494.5	39.7	361	T06708	cysteine proteinase
4	493.5	39.6	462	J08715	broccoli inducible
5	491	39.4	427	S57776	cysteine proteinase
6	490.5	39.4	456	T06416	cysteine proteinase
7	488.5	39.2	374	T03941	cysteine proteinase
8	485.5	39.0	378	S47434	cysteine proteinase
9	484	38.9	334	KHMSL	cathepsin L (EC 3.
10	483	38.8	455	T12041	cysteine proteinase
11	481	38.6	362	S12581	cysteine proteinase
12	480	38.6	472	C05752	cysteine proteinase
13	479.5	38.5	428	CQ1121	cysteine proteinase
14	479	38.5	334	KHRTL	cathepsin L (EC 3.
15	477	38.3	333	KHHL	cathepsin L (EC 3.
16	475.5	38.2	364	T47430	cysteine proteinase
17	474.5	38.1	364	T12039	cysteine proteinase
18	472.5	38.0	346	TA0154	cysteine proteinase
19	470.5	37.8	360	S57777	cysteine proteinase
20	470.5	37.9	458	KPE20A	oryzain (EC 3.4.22
21	469	37.7	360	T08122	cysteine endopepti
22	469	37.7	471	KH2208	oryzain (EC 3.4.22
23	467	37.5	318	KHHL	cathepsin L (EC 3.
24	466.5	37.5	343	D86198	cysteine proteinase
25	465.5	37.4	334	A50195	cathepsin L (EC 3.
26	465.5	37.4	355	T06122	cysteine proteinase
27	464.5	37.3	341	A42482	cathepsin S (EC 3.
28	462.5	37.1	465	T04208	cysteine proteinase
29	460.5	37.0	217	E15841	cathepsin S (EC 3.

ALIGNMENTS

RESULT 1

JC4848

cysteine proteinase (EC 3.4.22.) Douglas fir

N:Alternate names: pseudotrypsin

C:Species: Pseudotsuga menziesii (Douglas fir)

C:Date: 15-Aug-1996, #sequence_revision:15-out 1996 #text_change:64-Feb-2000

C:Accession: JC4848

R:Tranbarger, T.J.; Misra, S.

Cite: 172, 221-226, 1996

A:Title: Structure and expression of a developmentally regulated cDNA encoding a cyst

A:Reference number: JC4848, MUID:96269408

A:Accession: JC4848

A:Molecule type: mRNA

A:Residues: 1-454 <TR>

A:Cross-references: Ch.041.002, NID-91208548, FIDN:AA03455.1, FID:31238549

A:Note: It is uncertain whether Met-1, Met-15 or Met-41 is the initiator

C:Summary: This enzyme catalyzes the initial stages of storage protein mobilization d

C:Superfamily: papain

C:Keywords: cysteine proteinase, hydrolase

F:156,292,312/Active site: Cys, His, Asn #status predicted

Query Match 40.9%; Score 509; DB 2; Length 454;

Best local Similarity 48.6%; Pred. No 1 9e-37;

Matches 109, Conservative 32, Mismatches 74, Indels 8, Gaps 6;

QY 2 INSLHWVAGGVNT:FNKNKNGSSWAFSAVAALSSGATINELGFLSELSEQFVDCSKON 61

LE 134 PESLWREKGAIVAKNNSGSSWAFSAVAALSSGATINELGFLSELSEQFVDCSKON 190

QY 62 QNFQVWETIMGLAPYAIKNKYLDINLTPYPAEAFKFKDMSEFNFTIPVFKAYKYVPPR 121

DE 191 YNCGTNGTLMVYAEFTISNGTISFTQYPPANNSS-PAVPPNAHVVTITVETVFEN 249

QY 122 RNALFALAKYDLSVAIAQAGTIPYFYSVVLAPLQIKVNFVYVAYMEEDINKE 181

DE 250 DEKSFKAAND-PLISVAIFASAPAPYESSVPSNNGTIDIDGATVIVY--GSFSGID 306

QY 183 YVLVNSWELAWJEGYIFLA-LLISSEFFETGGIVHVVIV 201

DE 307 YVLVNSWNSWEGEPKIKLNLEGASIGMGIAEASYPV 348

RESULT 2

EM5435

cysteine proteinase (EC 3.4.22.) Pseudotsuga [sequenced] Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 18-Feb-2001, #sequence_revision:18-Feb-2001 #text_change:23-Mar-2001

C:Accession: EM5435

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

Nature 400, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

[illegible]

A:Accession: S00424
 A:Molecule type: protein
 A:Residues: 114-147, 'P', 148-220, 'X', 222-267, 'N', 269-288, 292-333 <NIT>
 R: Machleidt, W.; Ritonja, A.; Popovic, I.; Kotnik, M.; Brzin, J.; Turk, V.; Machleidt, W.
 In: Cysteine proteinases and their inhibitors, Turk, V., ed., pp.3-18, Walter de Gruyter,
 A:Title: Human cathepsins B, H and L: characterization by amino acid sequences and some
 A:Reference number: A27011
 A:Accession: B27011
 A:Molecule type: protein
 A:Residues: X, 115-129, 'M', 131-133, 'E', 135-141, 292-307, 'TP', 310-333 <MA2>
 R: Mason, R.W.; Walker, J.E.; Northrop, F.D.
 Biochem. J. 240, 373-377, 1986
 A:Title: The N-terminal amino acid sequences of the heavy and light chains of human cathepsin B
 A:Reference number: A26069; MUID:87127952
 A:Accession: A26069
 A:Molecule type: protein
 A:Residues: 114-147, 'P', 149-152, 'Y', 292-333 <MAS>
 R: Smith, S.M.; Gottesman, M.M.
 J. Biol. Chem. 264, 20487-20495, 1989
 A:Title: Activity and deletion analysis of recombinant human cathepsin B expressed in Escherichia coli
 A:Reference number: A32683; MUID:90062183
 A:Contents: annotation
 C:Genetics:
 A:Gene: GBA:CTSL
 A:Cross-references: GDB:119824; OMIM:116880
 A:Map position: 9q22.1-9q22.2
 A:Introns: 42, 83, 85, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132
 C:Function: heterodimer of disulfide linked chains produced from a single chain precursor
 A:Description: catalyzes hydrolysis of peptide bonds in proteins
 A:Pathway: intracellular protein degradation
 A:Note: Important role in the lysosomal degradation of proteins
 C:Superfamily: papain
 C:Keywords: cysteine proteinase; glycoprotein; heterodimer; hydrolase; lysosome; protein
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-113/Domain: propeptide #status predicted <PRO>
 F:114-333/Product: cathepsin B #status experimental <MAT>
 F:114-288/Product: cathepsin B heavy chain #status experimental <CHC>
 F:292-333/Product: cathepsin B light chain #status experimental <LCH>
 F:135-178, 169-211, 269-322/Disulfide bonds: #status predicted
 F:138, 276, 390/Active sites: Cys, His, Asn #status predicted
 F:221/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.3%; Score 477; DB 1; Length 339;
 Best Local Similarity 43.6%; Pred. No. 9e-35;
 Matches 102; Conservative 42; Mismatches 72; Indels 28; Gaps 8;
 QY 2 PNSTINWEAGCVNPIRNQNGSGWAFSAVALEGATCAQTNRGLPISLSQQDFVDCSKON 61
 DB 115 PRSDWHPKGYVTPVKNQDQGSQWAFSATGALPQMPFKTGP-LISLSEQLNLDVCSGPQ 173
 QY 52 GNFGCGGTWGLAFQVAKKRYLCTNCGYFYFAEKKTCM PSICFNVIETIV 112
 DB 174 GNFGCGGLMDYAFQVQVINGGLDPSFSPYATSTESCKYKYSVANDI---GFVDI--- 228
 QY 113 KAYKYVEPPNAINAKTALAFYDPSVAIQADQTPPEAFYKSGV-PPAPCGTK-VNKGVVIV 170
 DB 229 -----PRQEKALMKAVATGPIVAIDAGHESFLEYKRGIVTFPDSSSDMDHGVLV 281
 QY 171 EYDMQ---EDIRNKYWI-VNKGWGLAWGKRYIKALHSCKKCTGGHIVEPVYVI 222
 DB 282 GYGFESTSDNNKYWLIVKNSWGEGMGYVKMA--KDRNHCGIASAASYPTV 333

Search completed: October 9, 2002, 08:23:05
 Job time: 57 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

GenCore version 4.5
Copyright (c) 1992 - 2000 GenCorp Inc

OM protein - protein search, using sw model

Run on: October 9, 2002, 06:01:06 Search time: 66.62 seconds
(without alignments)
90,567 million cell updates/sec

Title: US-09-598-062-6
Sequence: 1 PPSINWVFAVCNPIPNQK.....KZTGTGILFVWPVWVQSI 226

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493.5	39.6	462	1	RD21-ARATH
2	489.5	39.3	334	1	CSE2_HUMAN
3	484	38.9	334	1	CATL_MOUSE
4	481	38.6	362	1	CSE_VLGM0
5	480.5	38.6	217	1	CATL_PHEP
6	480	38.6	362	1	CYSP_PHAVU
7	479.5	38.5	328	1	CYS4_HFANA
8	479	38.5	344	1	CATL_PAT
9	477	38.3	333	1	CATL_HUMAN
10	475.5	38.2	334	1	CATL_BOVIN
11	472.5	38.0	346	1	CYSL_LYCES
12	470.5	37.8	360	1	CYSP_HMSF
13	470.5	37.8	458	1	ORVA_GPYSA
14	469	37.7	471	1	ORVB_ORVSA
15	465.5	37.4	344	1	CATL_PIG
16	464.5	37.3	331	1	CATL_HUMAN
17	462	37.1	218	1	CATL_CHICK
18	460.5	37.0	217	1	CATL_BOVIN
19	459	36.9	340	1	CATL_MOUSE
20	455.5	36.6	329	1	CATL_PABIT
21	454	36.5	321	1	CYS3_HOMAM
22	454.5	36.4	380	1	ATIN_ACTH
23	452.5	36.3	329	1	CATL_RAT
24	452.5	36.3	330	1	CATL_PIG
25	451.5	36.3	329	1	CATL_HUMAN
26	451.5	36.3	329	1	CATL_MACFA
27	446	35.8	345	1	ANAN_ANANU
28	442.5	35.5	371	1	CYSL_HRPV3
29	442.5	35.5	373	1	CYS2_HRPV1
30	442	35.5	376	1	CYS2_PHELI
31	439.5	35.3	323	1	CATL_MOUSE
32	438.5	35.2	343	1	PAP3_CARPA
33	436.5	35.1	221	1	SPIL_ZINOF

P02473 number of
P14080 carica papa
P15342 rattus norv
Q02765 rattus norv
Q95029 drosophila
P14518 ananas como
Q06964 entamoeba h
Q01957 entamoeba h
P13277 homarus ame
Q97014 mus muscula
P05994 carica papa
Q01958 entamoeba h

ALIGNMENTS

RESULT 1
RD21-ARATH
ID RD21-ARATH STANDARD: PRT; 462 AA.
AC P43297;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cysteine proteinase RD21A precursor (EC 3.4.22.-).
GN P421A OR A11547128 OR P2019_31.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurossid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=93314560; PubMed=83255504;
RA Koizumi M., Yamaguchi-Shinozaki K., Tsuji H., Shinozaki K.;
RT "Structure and expression of two genes that encode distinct drought-
inducible cysteine proteinases in Arabidopsis thaliana.";
RL Gene 129.175-182(1993).
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=9116719; PubMed=11130712;
RA White G., Alonso J., Altieri J., Araujo F., Bowman C.I., Brooks S.Y.,
PA Pfeiffer P., Chan A., Chen C., Chen H., Cheuk R.F., Chin C.W.,
EA Chao M.K., Guan L., Gonyea A.B., Gonyea A.F., Greasy T.H., Dewar K.,
PA Dunn P., Elqui P., Feldblyum T.V., Feng J.-D., Feng B., Fujii C.Y.,
PA Gill J.F., Goldsmith A.B., Haas B., Hansen N.F., Hughes B., Huizar L.,
PA Hunter J.I., Jenkins J., Johnson-Ropes J., Khan S., Khaykin E.,
EA Kim G.J., Koc H., Krenetzka L., Korte D.B., Kwan A., Lam B.,
EA Landin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
EA Lin X., Liu S.X., Liu J.A., Lures J.S., Maiti P., Marziali A.,
EA Meltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
EA Paj G., Petersen S., Pham P.K., Phipps M., Phipps T., Phipps D.,
EA Sakano H., Salberg S.I., Schwarz T.P., Shino P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
EA Ziegler L., Van Aken S., Vaynsberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nucleic Acids Res. 28(2000).

-I- INDUCTION: BY HIGH SALT CONDITIONS.

-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
PAPAIN FAMILY OF THIOL PROTEASES.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
or send an email to license@ebi.ac.uk).

[illegible][illegible]

FT	NON_VER	1	1

NR InterPro; IPR006668; peptidase_C1.

DR MEROPS: C01.010; --
 DR InterPro: IPR000866; FR_Tarpet
 DR InterPro: IPR000668; Peptidase_C1
 DR InterPro: IPR002132; Ribosomal_L5
 DR Pfam: PF00112; Thiolprot_act_site
 DR PRINTS: PR00705; PAPA1N
 DR PROSITE: PS00614; FR_TARPET; UNKNOWN_L1
 DR PROSITE: PS00358; RIBOS_MAL_L5; UNKNOWN_L1
 DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1
 DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 KW Hydrolase; Thiol protease
 SQ SEQUENCE 359 AA: 49632 MW: 63145D101A733R4F CRC64;

Query Match 40.9%, Score 509, DB 10, Length 370,
 Best local Similarity 40.6%, Pred. No. 5e 42,
 Matches 103; Conservative 32; Mismatches 78; Indels 8; Gaps 6;

QY 2 PSINWVFAAGVNPINPKNNSWAFSAVAALGCAICAGINRGLSLSGGAVGCSKON 61
 DB 130 PTVNPRKGGATVTKNGSGSWAFSAVAALGCAICAGINRGLSLSGGAVGCSKON 188
 QY 62 GHEGCGGTGMLAPYAIPRYLTNDYTPFAEKTMDSPCNYTEIFVKAYKYVEFF 121
 DB 189 S-GGNWIMVAFDFANWGLSSEDSVPYLAQKSGSSE--ANSAAVTHVYQVPPN 244
 QY 122 NINAKTALAKYPSVAIQAGTTPQFYKSCVFAFGKVNHEVIVVEYDMDENKE 181
 DB 245 NEAALMKAVANG-PVSAIEASGAFQYSGVPSGSGHGTETEDPSVAAGVGMPPG-GKK 402
 QY 182 YWLVNSWCEAMGKGYIKLAII SCKKTCGILVEIVYV 221
 DB 303 YWLVNSWCEAMGKGYIKLAII SCKKTCGILVEIVYV 343

RESULT 5
 Q40922
 ID Q40922 PRELIMINARY: PRT: 454 AA.
 AC Q40922; Q40919;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PSEUDOTZAIN (CYSTEINE PROTEASE).
 GN PM33CYP.
 OS Pseudotsuga mertensiana (Coniferales).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pseudotsuga.
 OX NCBI_TaxID: 4357;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=[MIRB] FRANCO;
 RX MEDLINE=96265408; PubMed 8618207.
 RA Tranbarger T.J., Misra S.;
 RA "Structure and expression of a developmentally regulated cDNA encoding
 RT a cysteine protease (pseudotsain) from Douglas fir";
 RI Gene 172:221-226 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=[MIRB] FRANCO;
 RA Tranbarger T.J., Misra S.;
 RL Physiol. Plantarum 0:0-0(0).
 RN [3]
 RP SEQUENCE OF 247-347 FROM N.A.
 RC TISSUE WHOLE SEEDLING;
 RA Tranbarger T.J., Misra S.;
 RL Submitted (MAP-1996) to the EMBL/GenBank/DBI databases.
 DR EMBL: 041902; AAC49455.1; --
 DR EMBL: 249765; CAAB9835.1; --
 DR HSP: P07711; 1CJL.
 DR MEROPS: C01.029;
 DR InterPro: IPR000118; Granulin.
 FT NON_TER

DR InterPro: IPR000668; Peptidase_C1
 DR InterPro: IPR001211; PLP_A2
 DR InterPro: IPR000169; Thiolprot_act_site
 DR Pfam: PF00396; granulin; 1
 DR Pfam: PF00112; Peptidase_C1; 1
 DR PRINTS: PR00705; PAPA1N
 DR SMART: SM00277; GRAN; 1
 DR PROSITE: PS00118; PA2_HIS; UNKNOWN_L1
 DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1
 DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 KW Hydrolase; Protease; Thiol protease
 SQ SEQUENCE 454 AA: 49719 MW: 482193R8P5A6D23F CRC64;

Query Match 40.9%, Score 509, DB 10, Length 454,
 Best local Similarity 40.6%, Pred. No. 6,7e 42;
 Matches 108; Conservative 32; Mismatches 74; Indels 8; Gaps 6;

QY 2 PSINWVFAAGVNPINPKNNSWAFSAVAALGCAICAGINRGLSLSGGAVGCSKON 61
 DB 133 PTVNPRKGGATVTKNGSGSWAFSAVAALGCAICAGINRGLSLSGGAVGCSKON 190
 QY 62 GHEGCGGTGMLAPYAIPRYLTNDYTPFAEKTMDSPCNYTEIFVKAYKYVEFF 121
 DB 191 YWQCGNGGIMYAFQPIITNSGLSSEDSVPYKANNSSC-DAYRKNNAHWITIDYEDVPEN 249
 QY 122 NINAKTALAKYPSVAIQAGTTPQFYKSCVFAFGKVNHEVIVVEYDMDENKE 181
 DB 250 DEKSLKKAANA-PISVAIEASGAFQYSGVPSGSGHGTETEDPSVAAGVGMPPG-GKK 306
 QY 182 YWLVNSWCEAMGKGYIKLAII SCKKTCGILVEIVYV 221
 DB 307 YWLVNSWCEAMGKGYIKLAII SCKKTCGILVEIVYV 348

RESULT 6
 Q9XF79
 ID Q9XF79 PRELIMINARY: PRT: 370 AA.
 AC Q9XF79;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PAPA1N-LIKE CYSTEINE PROTEASE (FRAGMENT).
 GN PRT15.
 OS Sandersonia aurantiaca (Christmas bells) (Chinese-lantern lily).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Malvaceae; Malvaceae; Liliaceae; Liliaceae;
 OC Sandersonia.
 OX NCBI_TaxID: 61864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eason J.R., Bucknell T.T.;
 RA "Programmed cell death during flower senescence: Isolation and
 RT characterization of cysteine proteases from Sandersonia aurantiaca";
 RI Submitted (MAP-1999) to the EMBL/GenBank/DBI databases.
 DR EMBL: AF133838; AAC28476.1; --
 DR HSP: P06785; 2AOT.
 DR MEROPS: C01.029; --
 DR InterPro: IPR000118; Granulin.
 DR InterPro: IPR000668; Peptidase_C1
 DR InterPro: IPR001211; PLP_A2
 DR InterPro: IPR000169; Thiolprot_act_site
 DR Pfam: PF00396; granulin; 1
 DR Pfam: PF00112; Peptidase_C1; 1
 DR PRINTS: PR00705; PAPA1N
 DR SMART: SM00277; GRAN; 1
 DR PROSITE: PS00118; PA2_HIS; UNKNOWN_L1
 DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1
 DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1
 KW Hydrolase; Protease; Thiol protease.
 FT NON_TER

*Isolation of a cysteine protease cDNA clone (SFCF1) from senescent
 RI leaves of sweet potato (*Ipomoea batatas* cv. 'Batong 62').
 KL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases
 DR EMBL: AF242372; AAK23968.1; ...
 DR HSP: P07111; ICDL.

DR MEROPS: C01.01A; ...
 DR InterPro: IPR000169; Peptidase_C1.
 DR InterPro: IPR000169; Thiol_prot_act_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPA1N.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 DR PROSITE: PS00640; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 KW Hydrolyase; Protease; Thiol protease.
 SQ SEQUENCE 339 AA; 46716 MW; 23041AB5D5E389F5 CRC64;

Query Match 39.7%; Score 494.5; DB 10; Length 349;
 Best Local Similarity 45.9%; Pred. No. 1.2e-40;
 Matches 101; Conservative 45; Mismatches 79; Indels 5; Gaps 5;
 QY 2 PNSTNWKVTAQVAG IFAQ PPTSTWAFSAVAALPGATATALINFTVLELSEKEVLSKCN 61
 DE 122 PTVDMWTGAVTPVKDNGCGTQSWAFSAVAAMEITKLSQNLISSEGLVMDCKG 180
 QY 62 CNFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 121
 DE 181 TDQGTGEGGMDQAFSPINNKGTITSNFYDITSSCKSKSSN SAAKISGYEDVIAN 239
 QY 122 NINAKTALAKYGPISVAIDALQTPPEFYKSSVFLAPQIKVNHGVLVEYMDSDTKE 181
 DE 240 SESALEKAVANG-PVSVVAIDAGSSDPQYSSGVFTGCGTLDHGVTAVGYGIAPYISK 297
 QY 192 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 220
 DE 298 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 327

RESULT 10
 Q9STI4 PRELIMINARY; PRI: 361 AA.
 ID Q9STI4
 AC Q9STI4
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CYS-TERE ENDOPEPTIDASE-LIKE PROTEIN.
 GN T29H11.140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Euprotiophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID: 3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Choisme N., Robert C., Brothier P., Wincker P., Cottolico L.,
 RA Artiguenave F., Sarric W., Weissbach J., Mores H.W., Paul S.,
 RA Benke K., Mayer R.F.N., Guelier P., Salancat M.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases
 [2]

QY 2 PNSTNWKVTAQVAG IFAQ PPTSTWAFSAVAALPGATATALINFTVLELSEKEVLSKCN 61
 DE 122 PTVDMWTGAVTPVKDNGCGTQSWAFSAVAAMEITKLSQNLISSEGLVMDCKG 180
 QY 62 CNFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 121
 DE 181 TDQGTGEGGMDQAFSPINNKGTITSNFYDITSSCKSKSSN SAAKISGYEDVIAN 239
 QY 122 NINAKTALAKYGPISVAIDALQTPPEFYKSSVFLAPQIKVNHGVLVEYMDSDTKE 181
 DE 240 SESALEKAVANG-PVSVVAIDAGSSDPQYSSGVFTGCGTLDHGVTAVGYGIAPYISK 297
 QY 192 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 220
 DE 298 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 327

DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 KW Hydrolyase; Thiol protease.
 SQ SEQUENCE 339 AA; 46716 MW; 23041AB5D5E389F5 CRC64;
 Query Match 39.7%; Score 494.5; DB 10; Length 361;
 Best Local Similarity 46.4%; Pred. No. 1.4e-40;
 Matches 101; Conservative 37; Mismatches 73; Indels 9; Gaps 7;
 QY 2 PNSTNWKVTAQVAG IFAQ PPTSTWAFSAVAALPGATATALINFTVLELSEKEVLSKCN 60
 DE 122 PTVDMWTGAVTPVKDNGCGTQSWAFSAVAAMEITKLSQNLISSEGLVMDCKG 187
 QY 61 NENFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 120
 DE 188 NENFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 244
 QY 121 PNTAKTALAKYGPISVAIDALQTPPEFYKSSVFLAPQIKVNHGVLVEYMDSDTKE 180
 DE 245 NENALILKAVANG-PVSVVAIDAGSSDPQYSSGVFTGCGTLDHGVTAVGYGIAPYISK 301
 QY 121 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 221
 DE 302 KYWIVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 343

RESULT 11
 Q94HX1 PRELIMINARY; PRI: 462 AA.
 ID Q94HX1
 AC Q94HX1
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P2919.31/FC319.31.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Euprotiophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID: 3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk P., Kim C.T., Koesema F., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
 RA Gollis-Smith A.B., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
 RA Kamiya A., Katlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Ouedera C.S., Palm C.J.,
 RA Pham P.K., Quach H.T., Sakurai T., Satou M., Saki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Eckel J.E.,
 RA Arabidopsis cDNA clones.
 RL Submitted (JUN 2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AY039606; AAK2661.1; ...
 DR EMBL: AY039606; AAK2661.1; ...
 SQ SEQUENCE 462 AA; 57957 MW; 9C-63AC1F051A22A CRC64;

Query Match 39.6%; Score 493.5; DB 10; Length 462;
 Best Local Similarity 44.4%; Pred. No. 2.3e-40;
 Matches 99; Conservative 40; Mismatches 70; Indels 7; Gaps 6;
 QY 2 PNSTNWKVTAQVAG IFAQ PPTSTWAFSAVAALPGATATALINFTVLELSEKEVLSKCN 61
 DE 138 PNSTNWKVTAQVAG IFAQ PPTSTWAFSAVAALPGATATALINFTVLELSEKEVLSKCN 195
 QY 62 CNFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 121
 DE 196 YNENFQCTGATMELAFQVAIKKKYLCINDYPPFAEEKTCMESPEYIIPVKAYKYVEPP 254
 QY 122 NINAKTALAKYGPISVAIDALQTPPEFYKSSVFLAPQIKVNHGVLVEYMDSDTKE 181
 DE 255 SEESLKAVANG-PVSVVAIDAGSSDPQYSSGVFTGCGTLDHGVTAVGYGIAPYISK 311
 QY 192 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 223
 DE 311 YWLVNFWNEWLEAWTEKYLKLAHSCK-KTICGILLVEFYVPV 311


```

RN [2]
KP SEQUENCE FROM N.A.
KA HARRAK H., Tobacco23doh Z.
KI "Isolation and characterization of a drought-induced gene (tdi-65)
KR encoding a cysteine protease in Lycopersicon esculentum.";
KL Submitted (Jul 1994) to the EMBL/GenBank/DDBJ databases
DR EMBL: AJ004137; CA005894.1; -;
DR EMBL: AF172856; AAD48496.1; -;
DR HSP: P00785; 2ACT;
DR MEROPS: C01.029; -;
DR InterPro: IPR000118; Granulin
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR001211; PLP_A2.
DR InterPro: IPR000169; Thiolprot_act_site.
DR Pfam: PF00396; granulin; 1.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: P00705; PAPAIN.
DR SMART: SM00277; GRAN; 1.
DR PROSITE: PS00118; PA2_HIS; UKK00WL_1.
DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1.
DR PROSITE: PS00149; THIOI_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UKK00WN_1.
KW Hydrolase, Protease, Thiol protease, 466 AA; FEIE468967508BC7 CRC64;
SQ SEQUENCE 466 AA; 51141 MW; 1541446967508BC7 CRC64;

Query Match 35.44; Score 430.5; DR 10; Length 466;
Best Local Similarity 44.38; Pred. No. 4.7e-40;
Matches 98; Conservative 40; Mismatches 76; Indels 7; Gaps 6;

QY 2 DNSINWVEAGCVNFIENKNGSCWAFSAVAALLEGATCACTNPGLPSISFQOPVMSKQN 61
DB 139 PESIDWREKGVIVGVKINGSCGSAFSAVAAMESINAIVTG-NLISLSEQELVDCDR-S 196
QY 62 CNFQCTGCTMELAFQVAIKNKYLCTINDYVFAEKEKTCMSFCENYIELPVKAKYVPR 121
DB 197 YNEGCDGIMQYAFEVIRNMGIDTEEDYPYKRNQVC-DQFKKAKVVKYKIDSYEDVPVN 255
QY 122 NINAKTKALAKGFIKVAICAGTPEQFYKSGVFAPQGTKVNHGVLVLEVEYEMEDTNKE 181
DB 256 NEKALQKAVA-HQPVSLAFACGPNQHYKSGIETKQGVTAHVHGVVIACV--GTENQMD 312
QY 182 YWIVRNSWGEAWGEKGYIKALH-SGKKSTGGILVEPVYPV 231
DB 313 YWIVRNSWGANWGEKGYIKALH-SGKKSTGGILVEPVYPV 353

RESULT 15
ID 024147 PRELIMINARY; PRT; 374 AA.
AC 024147
DT 01-JAN-1998 (FEBBLrel. 05, Created)
DT 01-JAN-1998 (FEBBLrel. 05, Last sequence update)
DT 01-DEC-2001 (FEBBLrel. 19, Last annotation update)
DE CYSTEINE PROTEINASE PRECURSOR.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Becker G., Muentz K.;
RT "cDNA cloning of a CPR1-homologous proteinase from germinating tobacco
RT seeds.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: Z99174; CAB16317.1; -;
DR HSP: P00785; 2ACT;
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000169; Thiolprot_act_site.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: P00705; PAPAIN.
DR PROSITE: PS00640; THIOI_PROTEASE_ASN; 1.

```

```

DR PROSITE: PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; 1.
KW Hydrolase; Signal; Thiol protease.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 139 374 CYSTEINE PROTEINASE.
SQ SEQUENCE 374 AA; 42307 MW; 1541446967508BC7 CRC64;

Query Match 39.28; Score 488.5; DR 10; Length 374;
Best Local Similarity 44.98; Pred. No. 5.5e-40;
Matches 102; Conservative 37; Mismatches 79; Indels 9; Gaps 7;

QY 2 DNSINWVEAGCVNFIENKNGSCWAFSAVAALLEGATCACTNPGLPSISFQOPVMSKQN 61
DB 141 PHSVIMPKKFCVAPIKNGSCGSAFSAVAAGELNIVIGF-MHLSFQELVDCDRVQ 199
QY 62 CNFQCTGCTMELAFQVAIKNKYLCTINDYVFAEKEKTCMSFCENYIELPVKAKYVPR 121
DB 200 -NSGCDGIMQYAFEVIRNMGIDTEEDYPYKRNQVC-DQFKKAKVVKYKIDSYEDVPVN 256
QY 122 NINAKTKALAKGFIKVAICAGTPEQFYKSGVFAPQGTKVNHGVLVLEVEYEMEDTNKE 181
DB 257 NEKALQKAVA-HQPVSLAFACGPNQHYKSGIETKQGVTAHVHGVVIACV--GTENQMD 313
QY 182 YWIVRNSWGEAWGEKGYIKALH-SGKKSTGGILVEPVYVINGSI 236
DB 314 YWIVRNSWGEAWGEKGYIKALH-SGKKSTGGILVEPVYVINGSI 360

```

Search completed Oct-09 9, 2002, 08:59:17
Job time: 2149 sec

[illegible][illegible]

Query Match	59.5%	50.0%	49.0%	48.0%	46.0%	44.0%
Best Local Similarity	45.6%	45.0%	44.0%	43.0%	41.0%	39.0%

Do 115 PRVWPFKQYVTHVKNQCGGSCWAFSATGALFCOMFRTOR-11STSPQNI VDCSGPQ 173
QY 62 GNEQTCGIMCLAFQYAIENKXICINQYFYFAEKTOM-----USPOENYIEIPV 112
Db 174 GNEGCRGCLQYALQYVQCMCHITFESYIEATFESYINREYVARET GIVD1 228
QY 118 KAKYVETFRNAINALETALAFYDTSVAIQM-TEFFQYFESGV-ELAP-NTK-VNHQVILV 170
Db 229 -----PKZKALMRAVAIVGHSVAUAGHESFLFYREGYERFDASSEDMDRQVAV 281
QY 171 FYMDA--EDTNKYVWLVPNSWGEAMBEKQYTKIALHSCKFGTQCHLVEPVIWI 222
Db 282 GYCHPUSHSORRNYWLVZKSWGEWEMQGVKNA KORENHCJIASAASYPTV 333

Search completed: October 9, 2002, 08:25:01
Job time: 153 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000


```

AAW27446
ID AAW27446 standard; Protein: 374 AA.
AC AAW27446;
XX
DT 20-APR-1998 (first entry)
DE
DE Oil seed rape cysteine protease CDY5866.
XX
KW Promoter; oil seed rape; CDY5866;
KW Cysteine protease; disrupted protein; plant; tolerance; herbicide;
FW insect pest; fungal disease; improved yield; improved quality;
XX fertility control.
XX
XX Brassica napus
XX
XX W05745983 A2.
XX
XX 02-OCT-1997.
XX
XX 18-MAR-1997; 97WO-GR00729.
XX
XX 22-MAR-1996; 96GB-0006062.
XX
XX (ZENEC) ZENEC LTD.
XX
XX Greenland AJ, Jepson I, Thomas DRP;
XX
XX WPI: 1997-489646/45.
XX
XX N-PSDB: AAT90174.
XX
XX Cysteine protease promoter, obtained from oil seed rape, useful to
XX develop plants with improved agronomic characteristics
XX
XX disclosure; Fig 18; 137pp; English.
XX
XX The present sequence is an oil seed rape cysteine protease,
XX the promoter for which can restrict the expression of a disrupted
XX protein gene to a suitable stage of plant development, to provide
XX plants with novel agronomic features, e.g. tolerance to herbicides,
XX insect pests and fungal diseases, improved yield and/or quality of
XX harvested product, and novel mechanisms for the control of plant
XX fertility.
XX
XX Sequence 374 AA:

Query Match 38.6%; Score 481; DB 18; Length 374;
Best Local Similarity 45.9%; Pred. No. 4,3e-43;
Matches 102; Conservative 30; Mismatches 82; Indels 8; Gaps 6;

QY 2 PNSTINWVACGVNPIRNCQSCWAFSAVAALGATCACTNPGLPSISEQGFVDCSKQN 61
DQ 146 petyadarkkalkkkggqasawafsaalaglakktgqglisecpe-lvdcdr s 203
QY 62 GNFQCGGCTWHAFCQVAIAKPYLCINDDYTPFAEKTGM-----DSFENYIEIPV 121
DQ 204 pggqcmqladylqllkknltlctedjffriscjps gslpkasrltqgqadpke 202
QY 122 NIRAALAKAYGPIISVAIQAGTPPQFYKSGVFAFNTKVRNGSWVVEYEMBEDINKE 181
DQ 264 detalkravs yfprscslacayrvqyqysqlltqkctatldharrayq qseagid 319
QY 182 YMLVNSWGFAMWCKGVIFLA--LHSCKKGCTGTHLVEPVVPV 221
DQ 320 ymlvnswwtqwedgyfrrctalaraksqgqatqacasypp 361

RESULT 12
AAR28829
ID AAR28829 standard; Protein: 333 AA.
XX
AC AAR28829;
XX
DT 17-MAY-1993 (first entry)
DE
DE Human procathepsin L.
XX
XX Monoclonal antibody; cathepsin; diagnosis; cancer.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 94..107
XX /note= "claim 1, page 33"
XX Peptide 112..125
XX /note= "claim 1, page 33"
XX Peptide 282..295
XX /note= "claim 1, page 33"
XX Peptide 297..310
XX /note= "claim 1, page 33"
XX Peptide 308..321
XX /note= "claim 1, page 33"
XX
XX W09219756-A.
XX
XX 12-NOV-1992.
XX
XX 30 APR 1992; 92WO-JP00568.
XX
XX 02-MAY-1991; 91JP-0128274.
XX
XX (IDEK ) IDEMITSU KOSAN CO LTD.
XX
XX Irie S, Kawabata T, Miyayama T, Nakayama K, Yokoi E;
XX
XX WPI: 1992-398880/48.
XX
XX N-PSDB: AA031260.
XX
XX Monoclonal antibodies specific for human procathepsin L - and
XX five peptides derived from it, useful in diagnosis and
XX treatment of liver, lung, uterine and ovarian cancers
XX
XX Claim 1; Page 45; 48pp; Japanese
XX
XX Monoclonal antibodies specific for human procathepsin L and the
XX five peptides derived from it (indicated in the features) are claimed.
XX They are used in the diagnosis of liver, lung, uterine and ovarian
XX cancers.
XX
XX Sequence 333 AA:

Query Match 38.3%; Score 477; DB 13; Length 333;
Best Local Similarity 43.6%; Pred. No. 9.9e-43;
Matches 102; Conservative 32; Mismatches 72; Indels 28; Gaps 8;

QY 2 PNSTINWVACGVNPIRNCQSCWAFSAVAALGATCACTNPGLPSISEQGFVDCSKQN 61
DQ 115 pttvdkwrekayvlpvkqqqdcscwafsaalaglakktgqglisecpe-lvdcdr s 173
QY 62 GNFQCGGCTWHAFCQVAIAKPYLCINDDYTPFAEKTGM-----DSFENYIEIPV 112
DQ 174 gggcmqladylqyqyqadqgldscosyfycaatecsckypkysvadt---afvdi-- 228
QY 113 KAYKYVFPNTNIAKALAKAYGPIISVAIQAGTPPQFYKSGV-FDAPGCTK-VNHGVVLV 170
DQ 229 ptkaklmkavavtpqisvaldghesftfkyctfypdcscdmhdgvlvy 281
QY 171 LYEMD KSTNKYVILVENSQGFAMWCKGVIFLA--LHSCKKGCTGTHLVEPVVPV 222
DQ 282 gylfctesclndkylvklvswdeewqmggyvkaad-kdrlnhclgasasptv 333

RESULT 13
AAW47031

```


CC pathway and various developmental pathways in a foresty plant, by
 CC stable, participating in the present coding sequence into the genome
 CC of the foresty plant, where the coding sequence provides a PCD pathway
 CC that is not present in a native form of the foresty plant.
 XX
 SQ Sequence 272 AA

Query Match 17.7%, Score 158.5, DB 22, Length 272,
 Best Local Similarity 30.1%, Pred. No. 4, 26-08;
 Matches 43; Conservative 25; Mismatches 44; Indels 31; Gaps 4;

QY 34 KLENIATLAIPIVLVTVSVSLYTNNTSKITMVFVDPVGVVATQPY-PFSFEPRKK 92
 DB 29 qptltatltiv tsetletlprlqcehllkqesqmay 59
 QY 53 YPVVSEMEENLQELTETPRNNPFLI-INSQSFVY-LMPDQGLSEFTHMASFDTIR 151
 DB 60 harsykdckakr/yfthpntkafadkldayk/mtfddtthpntstsfykr 119
 QY 152 DSKUGKVFKSSKVSASLSEER 174
 DB 120 -----rstrvlsvdekpl 133

RESULT 6
 AAB85128
 ID AAB85128 standard; Protein; 345 AA.

XX AAB85128;
 DI 22-AUG-2001 (first entry)
 DE Carica cysteine proteinase.
 XX Cryptosporidium, infection; cryptogain; antiparasitic vaccine;
 KW immunization; cysteine proteinase.

XX Carica papaya.
 PN US6254869-B1.
 XX 03-JUL-2001.
 PD 27-MAR-1997; 970S-0827171.
 PF 27-MAR-1996; 960S-0014253.
 PR (REGC) UNIV CALIFORNIA.
 PA Petersen C, Huang J;
 PI WPI: 2001-388923/41.
 XX New vaccines and cryptogain protein for diagnosis and treatment of
 PT Cryptosporidium species infection

XX Example 2; Fig 4; 32pp; English.
 XX The invention relates to the uses and cryptogain protein for diagnosis
 CC and treatment of Cryptosporidium species infection. Cryptogain protein
 CC can be expressed by standard recombinant methods by the enzymes
 CC comprising the cryptogain protein or its fragments are useful for active
 CC immunization of animals and humans against Cryptosporidium infection, or
 CC for production of passive immune products in admixture with an adjuvant.
 CC Cryptogain can be used for prophylactic, therapeutic, diagnostic and
 CC detection purposes. The present sequence represents a cysteine proteinase
 CC from Carica, used in homology studies with the C. parvum cryptogain
 XX Sequence 345 AA;
 SQ

Query Match 17.6%, Score 157.5; DB 22; Length 345;

Best Local Similarity 32.5%, Pred. No. 7, 30-08;
 Matches 37; Conservative 26; Mismatches 46; Indels 5; Gaps 2;
 QY 40 IATLAIPIVLVTVSVSLYTNNTSKITMVFVDPVGVVATQPY-PATREYKSKFEPRKKYKV 96
 DB 1 mutipisiskilfvadelfvymals fgdslvqyspalltsteritqlfcswnlkuk; 58
 QY 97 YSSMPEPNORRPIYKONMNFIKTINSOGFSYVLEPMNFGDLSKEEFMARPTGYI 150
 DB 59 ykaidekkyfyeifkankk7dactokknas7wlglnvfnadsndefkexyqgsi 112

RESULT 7
 AAY83105
 ID AAY83105 standard; Protein; 311 AA.

XX AAY83105;
 AC 04-JUL-2000 (first entry)
 XX 710 proteinase from pineapple stem.
 DE Proteinase, pineapple, treatment, prevention; prophylaxis; tumour;
 KW solid tumour; non-solid tumour; virus induced tumour;
 KW immunostimulant; immunodeficiency; HIV;
 KW human immunodeficiency virus; malaria; malnutrition; trauma; burn;
 KW surgery; genetic abnormality; diabetes; nitric oxide; NO;
 KW antimicrobial.

XX Ananas comosus.
 OS W0209014253-A1.
 PN 16-MAR-2000.
 PD 24-AUG-1999; 97WC-GB02784.
 PR 02-SEP-1998; 98CR-0019138.
 PA (CORT-) CORTECS UK LTD.
 PI Mynott TL, Crossett B;
 XX WPI: 2000-266099/22.
 DE R-FSCB, AA293389.

XX stem bromelain like proteinases useful for the prevention and treatment
 PT of cancer and diseases or conditions which respond to increased nitric
 PT oxide production are encoded by nucleic acid sequences isolated from a
 PT pineapple stem -
 XX Claim 4; Page 58-59; 72pp; English.

XX The sequences encoding proteinases from pineapple stem have
 CC the same activity as a stem bromelain like proteinase, CCN2.
 CC The proteinases can be used in human or veterinary medicine and
 CC can be used for treating or preventing cancer, e.g. solid tumours;
 CC non-solid tumours and virus induced tumours. The proteinases can
 CC also be used as an immunostimulant or in the preparation of an
 CC immunostimulant to treat immunodeficiencies arising from
 CC malnutrition, infection (e.g. HIV and malaria), tumours (e.g.
 CC lymphoma, myeloma), trauma (e.g. lacerations, wounds, surgery) medical
 CC treatment (e.g. with drugs such as steroids, cyclosporin and
 CC cytochalasin), protein loss (e.g. in diarrhoea and burns),
 CC genetic abnormalities (e.g. in combined immunodeficiency patients
 CC lacking B or T cells), diabetes and old age. The proteinase can
 CC also be used in the treatment or prevention of diseases or
 CC conditions which respond to increased nitric oxide (NO) production,
 CC or as an antimicrobial agent or the component of an antimicrobial
 CC agent.

SQ Sequence 311 AA;

PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PP 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 15-JUN-1999; 990S-0139453.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139442.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140841.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142955.
PR 06-JUL-1999; 990S-0143390.
PR 08-JUL-1999; 990S-0143803.
PR 09-JUL-1999; 990S-0144290.
PR 12-JUL-1999; 990S-0144297.
PR 13-JUL-1999; 990S-0144342.
PR 14-JUL-1999; 990S-01443624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145216.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147402.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147203.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149829.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151338.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 22-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159429.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 18-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.


```

PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148419.
PR 12-AUG-1999; 990S-0148441.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149468.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151989.
PR 30-AUG-1999; 990S-0151903.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152463.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155189.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155559.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157763.
PR 06-OCT-1999; 990S-0157855.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159431.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159384.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161459.
PR 26-OCT-1999; 990S-0161460.
PR 26-OCT-1999; 990S-0161461.
PR 28-OCT-1999; 990S-0161929.
PR 28-OCT-1999; 990S-0161932.
PR 28-OCT-1999; 990S-0161933.
PR 29-OCT-1999; 990S-0162142.

Query Match: 17.3%; Score 154.5; DB 21; Length 358;
Best Local Similarity 33.0%; Prod No. 1 6e-07;
Matches 39; Conservative 19; Mismatches 43; Indels 15; Gaps 2;

QY 41 ATLIAPIVLVTVSLVLTNNISDKIDQVPGDYVPAIKPEYK-----SFE 87
DE 2 kktlssvvlvivaasaaaiq--fdespirmvdsdgirowesvslqgsthvlsfa 60
QY 88 PPKKRYKVVYSSMFEENQPEEYKQNNPFIKTNSQSPSVIFMNPFFDLSKEEF 142
DB 61 rfthrygkkyvcoemkifisiklenldirstakkglskyqvoqfadltwqef 115

RESULT 11
AAG47118
ID AAG47118 standard, Protein; 358 AA.
XX
AC AAG47118;
XX
DI 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59353.
XX
KW Protein identification, signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PI P1934405-A2.
XX
PD 06-SEP-2000.
XX
PF 25 FEB 2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 24-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129445.
PR 18-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130841.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132457.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.

```



```
db      3 4P+llsswv...l...gaaacacq-ldes:prrsdjlrveevsvlltqsrhlvsfa 60
QY     98 EFKKYYHKVYSMEERNPETLYQNMFITKTNSGSPSYVFMPNEFDLSKEEF 142
       : : : : : : : : : : : : : : : : : : : : : : : : : :
ob     61 rltthyakkygrveemkrslkngldirstnkkfysyklqvotadllwgef 115
```

APPENDIX 13

RESOL 12
AAW1954 I
10 AAW1954 I standard; Protein; 46+ AA.

AA AAW19541:

17-SEP-1997 (first entry)

XX Soybean thiol protease B3-alpha.

XX
KW
KW

XX (5) Glycine max.

For the purpose of this study, the following hypotheses were formulated:

Peptide

1. If $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, then $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$.

FT Protein 132.463

FETAL HEMI-PARALYSIS

JP09121870-A.

13-MAY-1997

28-1)F"-1995- 951P-035431

XX
30-APR-1995
95 PD-0245 376

96-10-1245
 96-10-1246
 96-10-1247
 96-10-1248

XX
PA (A.I.N.) A.I.N.(MOT) KKXX
 1007-314334730

DOI: 10.1002/polb.10001
WPI: 1999-3143307
N-PSDH: AA174041

XX New thiol protease and corresponding DNA sequence - used in the
PT preparation of the thiol protease
PT

xx
ps
claim 2; page 10-12; 17pp; Japanese.

The present sequence represents a novel thiol protease B3-alpha derived from denatured soybean cotyledon. A method for the production of thiol protease using *Escherichia coli* transformed with recombinant expression vectors containing a DNA sequence encoding thiol protease has been produced. The method can be used to prepare large amounts of thiol protease with high efficiency.

XX	Sequence	463 AA;
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Query Match:	17.09	Score	150	DB	18	Length	463
--------------	-------	-------	-----	----	----	--------	-----

Best Local Similarity 31.68; Pred. No. 4e-07;

Base local similarity 27, 28; 40–42;
Matches 36; Conservative 27; Mismatches 37; Indels 14; Gaps 3;

40 1ATL1AIFIVVTVSLYITNNFŠDK[DQFVPGJYVDPAL- - -REYFKSEEEFKKKYHK 95

α	β	γ	δ	ϵ	ζ	η	θ	ι	κ	λ	μ	ν	ξ	\omicron	π	ρ	σ	τ	υ	ϕ	χ	ψ	ω
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21																			

STRENGTHENING THE U.S. ECONOMY

QY 96 VYSSMEEHNQRFFIYKQNNMFIKTTNSQ-GFSYVLFMNFEGDLSKEEFMAHFG 148

LIBRARY

RESULT
AACS3797

AAG53787
 iD AAG53787 standard; Protein; 325 AA.

[illegible]

protein identification, signal transduction pathway, metabolic pathway, hybridisation assay, protein mapping, gene expression control, promoter, termination sequence.

Arabisopsis thalictroides.

EP103405-A2.

06-SEP-2000

03 VLURU=CHL(U)IC *UUIC=PLD-SC

2006-11-26 10:00

05-MAR-1999: 99JS-0123180;
05-MAR-1999: 99JS-0123649

23-MAR-1999; 9905-0125788.
25-MAR-1999; 9905-0125784.

29-MAY-1999; 9905-0128785;
01-MAY-1999; 9905-0128785;
01-MAY-1999; 9905-0128785;

06-APR-1999; 99US-0128234.

16-APR-1999; 99US-0129845.

21-APR-1999 09:05-013049

23-APR-1999; 99US-0130891.

30-APR-1999: 99US-0132048.

04-MAY-1999; 9515-0132484

06-MAY-1994: 0911Z-0182486

07-MAY-1999; 99US-0132863.

14-MAY-1999 0905-0134218

14-MAY-1999; 99JJS-0134221.

14 MAY 1993, 9905-0134770.
18-MAY-1994, 99112-0134768

19-MAY-1999; 9905-0134941.
20-MAY-1999; 9905-0135124

21-MAY-1999; 9905-0135353.
24-MAY-1999; 9905-0135679.

25-MAY-1999; 99JJS-0136021.
27-MAY-1999; 99JJS-0136397

28-MAY-1999: 99JIS-0136782.
01-JUN-1999: 00JIS-0137327.

03-JUN-1999; 99US-0137528;
04-JUN-1999; 99US-0137503

07-JUN-1999; 99US-0137724.

10-JUN-1999; 99US-0138540.

14-JUN-1994; 99US-0134114.

16-JUN-1999; 99US-013945}.

18-JUN-1999: 99US-0139454.

18-JUN-1999; 99US-0139456.

‘h366i-Nii-8i

18-JUN-1999; 99JUS-0139460.

18-JUN-1999: 0651-NIN-81 2946310-5066

18-JUN-1999; 99US-0139750.

21-JUN-1994 09:38:17

22 JUN 1999; 99US-0139697.
23-JUN-1999; 99US-0140353.

100

PP	23-JUN-1999	9905-0140953
PP	24-JUN-1999	9905-0140959
PP	25-JUN-1999	9905-0140823
PP	26-JUN-1999	9905-0140991
PP	27-JUN-1999	9905-0141287
PP	28-JUN-1999	9905-0141842
PP	29-JUN-1999	9905-0142154
PP	30-JUN-1999	9905-0142055
PP	01-JUL-1999	9905-0142390
PP	02-JUL-1999	9905-0142803
PP	03-JUL-1999	9905-0142920
PP	04-JUL-1999	9905-0142977
PP	05-JUL-1999	9905-0143542
PP	06-JUL-1999	9905-0143624
PP	07-JUL-1999	9905-0144005
PP	08-JUL-1999	9905-0144045
PP	09-JUL-1999	9905-0144086
PP	10-JUL-1999	9905-0144335
PP	11-JUL-1999	9905-0144333
PP	12-JUL-1999	9905-0144333
PP	13-JUL-1999	9905-0144333
PP	14-JUL-1999	9905-0144334
PP	15-JUL-1999	9905-0144352
PP	16-JUL-1999	9905-0144352
PP	17-JUL-1999	9905-0144352
PP	18-JUL-1999	9905-0144352
PP	19-JUL-1999	9905-0144352
PP	20-JUL-1999	9905-0144352
PP	21-JUL-1999	9905-0144352
PP	22-JUL-1999	9905-0144352
PP	23-JUL-1999	9905-0144352
PP	24-JUL-1999	9905-0144352
PP	25-JUL-1999	9905-0144352
PP	26-JUL-1999	9905-0144352
PP	27-JUL-1999	9905-0144352
PP	28-JUL-1999	9905-0144352
PP	29-JUL-1999	9905-0144352
PP	30-JUL-1999	9905-0144352
PP	31-JUL-1999	9905-0144352
PP	01-AUG-1999	9905-0144352
PP	02-AUG-1999	9905-0144352
PP	03-AUG-1999	9905-0144352
PP	04-AUG-1999	9905-0144352
PP	05-AUG-1999	9905-0144352
PP	06-AUG-1999	9905-0144352
PP	07-AUG-1999	9905-0144352
PP	08-AUG-1999	9905-0144352
PP	09-AUG-1999	9905-0144352
PP	10-AUG-1999	9905-0144352
PP	11-AUG-1999	9905-0144352
PP	12-AUG-1999	9905-0144352
PP	13-AUG-1999	9905-0144352
PP	14-AUG-1999	9905-0144352
PP	15-AUG-1999	9905-0144352
PP	16-AUG-1999	9905-0144352
PP	17-AUG-1999	9905-0144352
PP	18-AUG-1999	9905-0144352
PP	19-AUG-1999	9905-0144352
PP	20-AUG-1999	9905-0144352
PP	21-AUG-1999	9905-0144352
PP	22-AUG-1999	9905-0144352
PP	23-AUG-1999	9905-0144352
PP	24-AUG-1999	9905-0144352
PP	25-AUG-1999	9905-0144352
PP	26-AUG-1999	9905-0144352
PP	27-AUG-1999	9905-0144352
PP	28-AUG-1999	9905-0144352
PP	29-AUG-1999	9905-0144352
PP	30-AUG-1999	9905-0144352
PP	31-AUG-1999	9905-0144352



Genere version 4.5
Copyright (c) 1993 - 2000 Computer Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 08:23:24 : Search time 49.6 seconds
(without alignments)
414.178 Million cell updates/sec

Title: US-09-598-062-5

Perfect score: 86%

Sequence: 1 MDIGNNVEEHQEVISGPYIA.....DERVKKSRVSASESEEPV 175

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 28138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 28138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	173	19.3	364	2 T12039	cysteine proteinas
2	170	19.0	368	2 T47812	cysteine proteinas
3	169	18.9	254	2 T45630	cysteine proteinas
4	163	18.2	348	2 JN0633	caricain (EC 3.4.2
5	163	18.2	367	2 JN0634	caricain (EC 3.4.2
6	158.5	17.7	464	2 S24602	cysteine proteinas
7	158	17.6	457	2 T07851	ananain (EC 3.4.22
8	157.5	17.6	445	1 P07A	papain (EC 3.4.22
9	156	17.4	254	2 A96441	cysteine proteinas
10	155.5	17.4	324	2 T10518	fruit bromelain (E
11	155.5	17.4	352	2 T09760	fruit bromelain (E
12	155	17.3	441	2 T45839	chymopapain (EC 3.
13	155	17.3	352	2 T10501	probable cysteine
14	154.5	17.2	656	2 T10514	fruit bromelain (E
15	154.5	17.2	456	2 T07849	probable stem brom
16	154.5	17.0	345	2 T07839	ananain (EC 3.4.22
17	150	16.7	340	2 T10516	fruit bromelain (E
18	150	16.7	377	2 T47471	cysteine proteinas
19	150	16.7	455	2 T12041	cysteine proteinas
20	149.5	16.7	351	2 T10503	fruit bromelain (E
21	147	16.4	625	2 S19451	cysteine proteinas
22	142.5	15.9	474	2 T06276	benzothiadiazole i
23	142	15.8	442	2 S21773	cysteine proteinas
24	142	15.8	355	2 T06222	cysteine proteinas
25	142	15.8	357	2 T06529	cysteine proteinas
26	141.5	15.8	476	1 K06CF	pestatalk cathepsin
27	140.5	15.7	448	2 T04672	prolactin cysteine
28	140	15.6	426	2 S53027	cathepsin L (EC 3.
29	137.5	15.3	474	2 T03941	cysteine proteinas

30 137.5 15.3 454 2 J04848 cysteine proteinas
31 137 15.3 356 2 T10436 cathepsin-like pro
32 137 15.3 380 1 T06B actinidain (EC 3.4
33 135 15.1 356 2 S66348 cysteine proteinas
34 134.5 15.0 323 2 J25591 cysteine proteinas
35 134.5 15.0 441 2 A45565 cysteine proteinas
36 134 15.0 348 2 T09798 glycyl endopeptida
37 133.5 14.9 324 2 T10394 cathepsin orgyia
38 133 14.8 569 2 A45634 trypsin-like cystei
39 132 14.7 314 2 S68783 cathepsin L (EC 3.
40 131 14.6 324 2 S62735 cathepsin - Choris
41 131 14.6 331 2 D86413 cysteine proteinas
42 130.5 14.6 323 2 S62736 cathepsin-like cys
43 130 14.5 360 2 S59598 cysteine proteinas
44 128.5 14.3 346 2 C86413 cysteine proteinas
45 128 14.3 361 1 K06H alourain (EC 3.4.2

ALIGNMENTS

RESULT 1

T12039

Cysteine proteinase (EC 3.4.22.1) : precursor - kidney bean

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

C:Accession: T12039

R:Senyuk, V.; Becker, C.; Muenitz, K.

submitted to the EMBL Data Library, October 1997

A:Description: Isolation of cDNA clone encoding cysteine proteinase (CPL) from a coty

A:Reference number: Z17385

A:Accession: T12039

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-364 <SEN>

A:Cross references: EMBL:Z99952

A:Experimental source: Cultivar Mojavia6, cotyledon, clone cp6a

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F1:1-19/domain: signal sequence #status predicted <SIG>

F1:20-124/chain: propeptide #status predicted <PRO>

F1:125-364/full:1 cysteine proteinase #status predicted <MAF>

F:149,285,305/Active site: Cys, His, Asn #status predicted

Query Match 19.3% Score 173, EB 2, Length 364;
Best local similarity 27.3%, Pred. No. 2,60-06;
Matches 35, Conservative 35, Mismatches 40, Indels 18, Gaps 2;

Cy 35 LKNITATLIAIPVLVTVSYLYTNNTSKIKDDVPGDYVDPATPEYKSFEEFKKKYH 94

DB 1 MFSLMFTLLLSLTSITSHATAMSTINYSFNFVMDM-----VEEWLVKRR 44

Cy 35 KYVSEMEINRGEFFVKKGMHRTNQGFSYVLEMHDFCDLSEKTFEMASITCYIKQSK 154

DB 45 KYVNGLEKRRGEFFVKKGMHRTNQGFSYVLEMHDFCDLSEKTFEMASITCYIKQSK 104

Cy 155 DERVEKKS 162

DB 105 --RRVVK 110

RESULT 2

S47312

Cysteine proteinase (EC 3.4.22.1) precursor - spring vetch

C:Species: Vicia sativa (spring vetch, tare)

C:Date: 06-Jun-1995 #sequence_revision 06-Jun-1995 #text_change 04-Feb-2000

C:Accession: S51817, S47312

R:Becker, C.; Fischer, J.; Nong, V.H.; Muenitz, K.

Plant Mol. Biol. 25, 1257-1262, 1994

A:Title: PCR cloning and expression analysis of cDNAs encoding cysteine proteinases f

A:Reference number: S51817, M010,95111103

A:Accession: S51817

[illegible][illegible]

A:Reference number: 416167
A:Accession: T07851
A:Status: Predicted, translated from CD, TMRI/DOB
A:McVee type: mRNA
A:Residues: 1-357 RGR
A:Cross-references: EMBL:AJ009830; PIDN:CAA08861.1
C:Genetics:
A:Gene: anll
C:Superfamily: papain
C:Keywords: Cysteine proteinase; hydrolase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-32/Product: anacain Anll #status predicted <MAT>
F:147-281-302/Active site: Cys, His, Asn #status predicted

Query Match 17.6% Score 158; DP 2; Length 357;
Best Local Similarity 32.9%; Prod No 3 9e-05;
Matches 35; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

Q7 42 LIAIDIVAVVSVSYITNTYDQIDREVPVGVYDPAIVPKSPFFKKYHKVYSSMEE 192
Db 7 LVELFLFCV--HWASISAASR-----DEPSDMMKRFPEWMAEYGRVYKONDE 53
QY 103 ENCRPEIVKQNMFKITNSQGFVSVLENNFEGLSKKEEPMARETG 148
Db 54 KMEFPEIQKKNVHIFTNSNGNSYITGINSOPTDMNFEFVAQYTG 100

RESULT 8
PPPA
papain (P0 2.4.22.2) precursor papaya
C:Species: Carica papaya (papaya)
C:Date: 24 Apr 1994 sequence revision 31-Mar-1992 #extl_change 18-Jun-1999
C:Accession A26466; A00974; A37903
R:Colgan, L.W.; Colgan, V.M.; Dibel, L.C.
Gene 48, 219-227, 1986
A:Title: Cloning and sequencing of papain-encoding cDNA.
A:Reference number: A26466; MUID:87181525
A:Accession: A26466
A:Molecule type: mRNA
A:Residues: 1-345 <OH>
A:Cross-references: GH:M15205; NID:q167399; PIDD:AA02850 1; PIDD:q167391
R:Mitchell, R.F.J.; Chaik-6, I.M.; Smith, F.L.
J. Biol. Chem. 245, 3485-3492, 1970
A:Title: The complete amino acid sequence of papain: Additions and corrections
A:Reference number: A92062; MUID:71007899
A:Accession: A00974
A:Molecule type: protein
A:Residues: 131-157, 279-281, 283-288, 291-292, 294-297, 299, 301, 303-307, 309, 311-312, 314-315, 317-318, 320-321, 323-324, 326-327, 329, 331-332, 334-335, 337-338, 340-341, 343-344, 346-347, 349-350, 352-353, 355-356, 358-359, 361-362, 364-365, 367-368, 370-371, 373-374, 376-377, 379-380, 382-383, 385-386, 388-389, 391-392, 394-395, 397-398, 400-401, 403-404, 406-407, 409-410, 412-413, 415-416, 418-419, 421-422, 424-425, 427-428, 430-431, 433-434, 436-437, 439-440, 442-443, 445-446, 448-449, 451-452, 454-455, 457-458, 460-461, 463-464, 466-467, 469-470, 472-473, 475-476, 478-479, 481-482, 484-485, 487-488, 490-491, 493-494, 496-497, 499-500, 502-503, 505-506, 508-509, 511-512, 514-515, 517-518, 520-521, 523-524, 526-527, 529-530, 532-533, 535-536, 538-539, 541-542, 544-545, 547-548, 550-551, 553-554, 556-557, 559-560, 562-563, 565-566, 568-569, 571-572, 574-575, 577-578, 580-581, 583-584, 586-587, 589-590, 592-593, 595-596, 598-599, 601-602, 604-605, 607-608, 610-611, 613-614, 616-617, 619-620, 622-623, 625-626, 628-629, 631-632, 634-635, 637-638, 640-641, 643-644, 646-647, 649-650, 652-653, 655-656, 658-659, 661-662, 664-665, 667-668, 670-671, 673-674, 676-677, 679-680, 682-683, 685-686, 688-689, 691-692, 694-695, 697-698, 700-701, 703-704, 706-707, 709-710, 712-713, 715-716, 718-719, 721-722, 724-725, 727-728, 730-731, 733-734, 736-737, 739-740, 742-743, 745-746, 748-749, 751-752, 754-755, 757-758, 760-761, 763-764, 766-767, 769-770, 772-773, 775-776, 778-779, 781-782, 784-785, 787-788, 790-791, 793-794, 796-797, 799-800, 802-803, 805-806, 808-809, 811-812, 814-815, 817-818, 820-821, 823-824, 826-827, 829-830, 832-833, 835-836, 838-839, 841-842, 844-845, 847-848, 850-851, 853-854, 856-857, 859-860, 862-863, 865-866, 868-869, 871-872, 874-875, 877-878, 880-881, 883-884, 886-887, 889-890, 892-893, 895-896, 898-899, 901-902, 904-905, 907-908, 910-911, 913-914, 916-917, 919-920, 922-923, 925-926, 928-929, 931-932, 934-935, 937-938, 940-941, 943-944, 946-947, 949-950, 952-953, 955-956, 958-959, 961-962, 964-965, 967-968, 970-971, 973-974, 976-977, 979-980, 982-983, 985-986, 988-989, 991-992, 994-995, 997-998, 1000-1001, 1003-1004, 1006-1007, 1009-1010, 1012-1013, 1015-1016, 1018-1019, 1021-1022, 1024-1025, 1027-1028, 1030-1031, 1033-1034, 1036-1037, 1039-1040, 1042-1043, 1045-1046, 1048-1049, 1051-1052, 1054-1055, 1057-1058, 1060-1061, 1063-1064, 1066-1067, 1069-1070, 1072-1073, 1075-1076, 1078-1079, 1081-1082, 1084-1085, 1087-1088, 1090-1091, 1093-1094, 1096-1097, 1099-1100, 1102-1103, 1105-1106, 1108-1109, 1111-1112, 1114-1115, 1117-1118, 1120-1121, 1123-1124, 1126-1127, 1129-1130, 1132-1133, 1135-1136, 1138-1139, 1141-1142, 1144-1145, 1147-1148, 1150-1151, 1153-1154, 1156-1157, 1159-1160, 1162-1163, 1165-1166, 1168-1169, 1171-1172, 1174-1175, 1177-1178, 1180-1181, 1183-1184, 1186-1187, 1189-1190, 1192-1193, 1195-1196, 1198-1199, 1201-1202, 1204-1205, 1207-1208, 1210-1211, 1213-1214, 1216-1217, 1219-1220, 1222-1223, 1225-1226, 1228-1229, 1231-1232, 1234-1235, 1237-1238, 1240-1241, 1243-1244, 1246-1247, 1249-1250, 1252-1253, 1255-1256, 1258-1259, 1261-1262, 1264-1265, 1267-1268, 1270-1271, 1273-1274, 1276-1277, 1279-1280, 1282-1283, 1285-1286, 1288-1289, 1291-1292, 1294-1295, 1297-1

C:Accession: T45839
 R:Kieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z24015

A:Accession: T45839
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-341 <RLE>
 A:Cross references: EMBL:AL132956
 A:Experimental source: cultivar Columbia; HAC clone F2K15
 C:Genetics:

A:Map position: 3

A:Introns: 149/1

A>Note: F2K15.200

C:Superfamily: papain

C:Keywords: cysteine proteinases; hydrolase

F:151/284/305/Active site: Cys, His, Asn #status predicted

Query Match 17.4% Score 155; DB 2; Length 341;

Best Local Similarity 28.9%; Pred. No. 6.3e-05;

Matches 37; Conservative 30; Mismatches 43; Indels 18; Gaps 4;

QY 46 IFVIVVTVSLYTRNGSKLSDVFGVGVVATPEYKSGFEELFKKTKHYKVVSSMEENQ 105

DB 5 VFELIALLSRSRTSVTS-----PQRIPEASAVE---KHFQMMSPRPVYSLSUSEKIS 54

QY 106 RFYFQRMHRIPTIN 22 IFVIVVIMRFTGLSPFFETMAFETLVYIFSKKQFVFKSSG 154

DB 55 RFELFNNAKEVSIANNINKYTIIDVNEFSOLDIREFKARYIGLVVPE-----GMIR 107

QY 165 VSASESPF 172

DB 108 ISTIPSEH 115

RESULT 13

T10501

fruit bromelain (EC 3.4.22.33) Fb1 precursor pineapple

C:Species: Ananas comosus (pineapple)

C>Date: 14 Jul 1999 #sequence_revision 14 Jul 1999 #text_change 20 Jun 2000

C:Accession: T10501

R:Muta, E.; Aramaki, H.; Takata, Y.; Kono, A.; Okamoto, Y.; Ota, S.

submitted to the EMBL Data Library, January 1993

A:Description: Cloning and sequencing of fruit bromelain

A:Reference number: Z17060

A:Accession: T10501

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-352 <NOT>

A:Cross references: EMBL:D14058

A:Experimental source: cv. N67-10; fruit; nearly mature

C:Function:

A:Description: hydrolases; proteins; cysteine proteinase

C:Superfamily: papain

C:Keywords: cysteine proteinases; hydrolase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-122/Domain: amino terminal propeptide #status predicted <PRO>

F:123-340/Domain: fruit bromelain #status predicted <MAL>

F:341-352/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:148/281/302/Active site: Cys, His, Asn #status predicted

Query Match 17.4% Score 155; DB 2; Length 352;

Best Local Similarity 32.1%; Pred. No. 6.5e-05;

Matches 35; Conservative 29; Mismatches 31; Indels 14; Gaps 3;

QY 43 LIALFVIVVTVSLYTRNGSKLSDVFGVGVVATPEYKSGFEELFKKTKHYKVVSSME 102

DB 7 LVFLFLFLAV---MWASPAASR-----DEFSDFMKKRPFEWMAEYGRVYKONDE 53

QY 103 DNGRFEVYKNNMFIFKTTAG-CUSAVIEMNFTGQSKDTMAFRTGY 150

DB 341FVIVVIMRFTGLSPFFETMAFETLVYIFSKKQFVFKSSG 154

DB 54 KMSEFGLKKNNVDB:IEFERNENSGVTLGINKFTIMTHNEFVAGVTCG 162

RESULT 14

T10514

probable stem bromelain (EC 3.4.22.32) precursor pineapple

N:Alternate names: FB38 precursor

C:Species: Ananas comosus (pineapple)

C>Date: 14 Jul 1999 #sequence_revision 14 Jul 1999 #text_change 20 Jun 2000

C:Accession: T10514

R:Muta, E.; Okamoto, Y.; Ota, S.

submitted to the EMBL Data Library, October 1994

A:Description: Cloning and sequencing of cysteine proteinases in Ananas comosus.

A:Reference number: Z17065

A:Accession: T10514

A:Status: preliminary, translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-356 <MUT>

A:Cross references: EMBL:D38532

A:Experimental source: cv. N67-10; nearly mature fruit

C:Function:

A:Description: hydrolases; proteins with broad specificity

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-122/Domain: amino terminal propeptide #status predicted <PRO>

F:123-338/Domain: stem bromelain #status predicted <MAT>

F:339-356/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:148/280/301/Active site: Cys, His, Asn #status predicted

Query Match 17.2% Score 154.5; DB 2; Length 356;

Best Local Similarity 42.1%; Pred. No. 7.2e-05;

Matches 32; Conservative 19; Mismatches 24; Indels 1; Gaps 1;

QY 84 KSPFEFKKTKHYKVVSSMEENORFEIYKNNMFIKTTNSQG-FSVVLENNFGLSKKEEF 142

DB 35 KPEEWMEVYGVYKNDKMKRPFQIFKNNVNHIEFNSPNNSYITGINTDTMTNNEF 94

QY 143 MAPFTGVTIKSDKDEP 158

DB 95 IAGYTCCISRPINTEP 110

RESULT 15

T07840

anain (EC 3.4.22.31) AN8 precursor - pineapple

C:Species: Ananas comosus (pineapple)

C>Date: 14 May 1999 #sequence_revision 14 May 1999 #text_change 20 Jun 2000

C:Accession: T07840

R:Robertson, C.E.; Goodenough, P.W.

submitted to the EMBL Data Library, July 1998

A:Description: Cloning and expression of anain gene from pineapple.

A:Reference number: Z16167

A:Accession: T07840

A:Status: preliminary, translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-356 <NOT>

A:Cross references: EMBL:AJ009829; F1UN:CAA0860.1

C:Genetics:

A:Gene: an8

C:Superfamily: papain

C:Keywords: cysteine proteinases; hydrolase

F:148/280/301/Active site: Cys, His, Asn #status predicted

Query Match 17.2% Score 154.5; DB 2; Length 356;

Best Local Similarity 42.1%; Pred. No. 7.2e-05;

Matches 32; Conservative 19; Mismatches 24; Indels 1; Gaps 1;

QY 84 KSPFEFKKTKHYKVVSSMEENORFEIYKNNMFIKTTNSQG-FSVVLENNFGLSKKEEF 142

DB 35 KPEEWMEVYGVYKNDKMKRPFQIFKNNVNHIEFNSPNNSYITGINTDTMTNNEF 94

CV 34 MAFEDYIKUSKUPR 100
DB 34 MAFEDYIKUSKUPR 100

Search completed: 00:00:00.000
34 Files (1) Sent

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGroup Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 09:01:05 ; Search time 96.62 seconds
(without alignments)
70 140 Million cell updates/sec

Title: US-09-598-062-5

Perfect score: 896

Sequence: 1 MDIGNNVEHQEYISGPYIA.....DERVEKSRVSASESSEEEV 175

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 105224 seqs, 48719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match %	Length	ID	Description
1	163	18.2	348	1 PAF3_CARPA	P10056 Carica papaya
2	157.5	17.6	345	1 PAF4_CARPA	P07784 Carica papaya
3	155.5	17.4	352	1 PAF2_CARPA	P14080 Carica papaya
4	152.5	17.0	345	1 ANAN_ANAGO	P08084 Ananas comosus
5	147.5	16.5	331	1 CATV_NPVES	P05944 Carica papaya
6	147.5	16.5	376	1 CATV_NPVES	P05944 Carica papaya
7	141.5	15.8	376	1 CATV_NPVES	P05944 Carica papaya
8	137	15.4	442	1 CATV_NPVES	P05944 Carica papaya
9	137	15.3	380	1 CATV_NPVES	P05944 Carica papaya
10	135	15.1	356	1 CATV_NPVES	P05944 Carica papaya
11	134.5	15.0	323	1 CATV_NPVES	P05944 Carica papaya
12	134.5	15.0	441	1 CATV_NPVES	P05944 Carica papaya
13	134	15.0	348	1 PAF4_CARPA	P14080 Carica papaya
14	134.5	14.9	324	1 CATV_NPVES	P05944 Carica papaya
15	133	14.8	569	1 CATV_NPVES	P05944 Carica papaya
16	131.5	14.7	446	1 CATV_NPVES	P05944 Carica papaya
17	131	14.6	324	1 CATV_NPVES	P05944 Carica papaya
18	130.5	14.6	323	1 CATV_NPVES	P05944 Carica papaya
19	129	14.5	460	1 CATV_NPVES	P05944 Carica papaya
20	129	14.4	324	1 CATV_NPVES	P05944 Carica papaya
21	128	14.3	462	1 CATV_NPVES	P05944 Carica papaya
22	127.5	14.2	422	1 CATV_NPVES	P05944 Carica papaya
23	127.5	14.2	505	1 CATV_NPVES	P05944 Carica papaya
24	127	14.2	443	1 CATV_NPVES	P05944 Carica papaya
25	126.5	14.1	360	1 CATV_NPVES	P05944 Carica papaya
26	126	14.1	444	1 CATV_NPVES	P05944 Carica papaya
27	126	14.1	462	1 CATV_NPVES	P05944 Carica papaya
28	125.5	14.0	323	1 CATV_NPVES	P05944 Carica papaya
29	125	14.0	324	1 CATV_NPVES	P05944 Carica papaya
30	124.5	13.8	362	1 CATV_NPVES	P05944 Carica papaya
31	122	13.6	423	1 CATV_NPVES	P05944 Carica papaya
32	120.5	13.4	462	1 CATV_NPVES	P05944 Carica papaya
33	120	13.4	421	1 CATV_NPVES	P05944 Carica papaya

RESULT ID	PAP3_CARPA	STANDARD:	PRT:	348 AA
AC	P10056:			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Caricain precursor (EC 3.4.22.30) (Papaya proteinase omega) (Papaya			
DE	proteinase III) (PPIII) (Papaya peptidase A).			
OS	Carica papaya (Papaya).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,			
OC	Spermatophyta, Malvales, Malvaceae, Caricaceae, Carica.			
OC	eurosid II, Brassicales; Caricaceae; Carica.			
OX	NCBI_TaxID:3649;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE:Leaf;			
PA	Baker K.C., Revell D.F., Cummings N.T., Collins M.E., Goodenough P.W.;			
PI	Submitted (MAY-1995) to the FMRB/GenBank/EMBL databases			
RN	[2]			
RP	SEQUENCE OF 133-348.			
PX	MEF1 INF-89105420; PubMed-4063283;			
PA	Dubois T., Kleinschmidt T., Schuck A.G., Iscoe Y., Braumiller G.;			
PI	"The thiol proteinases from the latex of Carica papaya L. II. The			
PI	primary structure of proteinase omega."			
RI	Biol. Chem. Hoppe-Seyler 269:741-754(1988).			
RN	[3]			
RP	SEQUENCE OF 237-348 FROM N.A.			
RC	TISSUE:Leaf;			
RA	Collins M.E., Revell D.F., Sumner I.G., Pickersgill R.W.,			
RA	Goodenough P.W.;			
PI	Submitted (FEB-1996) to the FMRB/GenBank/EMBL databases			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).			
PA	Pickersgill R.W., Fickelshaus P., Harris G.W., Goodenough P.W.,			
PI	"Determination of the structure of papaya protease omega."			
RL	Acta Crystallogr. B 47:766-771(1991).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF MUTANT G11 260.			
PX	MEF1 INF-96354872; PubMed-8769310;			
PA	Katerelos N.A., Taylor M.A.J., Scott M., Goodenough P.W.,			
RA	Pickersgill R.W.,			
RI	"Crystal structure of a caricain D158E mutant in complex with E-64."			
RL	FEBS Lett. 392:35-39(1996).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF P60-CARICAIN.			
PX	MEF1 INF-97094576; PubMed-8939744;			
PA	Greaves M.E., Taylor M.A., Scott M., Cummings N.J., Pickersgill R.W.,			
RA	Jenkins J.A.,			
RT	"The prosequence of precaricain forms an alpha-helical domain that			
RT	prevents access to the substrate-binding cleft."			
RI	Structure 4:1193-1203(1996).			
RN	[7]			
RP	3D-STRUCTURE MODELLING OF 133-348.			
PA	Blattl harpa S., Pal A., Bala A., Chakrabarty S., Banerjee A.,			
PI	"Theoretical model structure of a caricain D158E mutant in complex			

P35591 leishmania
P25494 hordeum vul
P25250 hordeum vul
P25776 oryza sativ
P22895 glycine max
P22497 theileria p
P12412 vigna mungo
P44296 arabidopsis
P25775 leishmania
P25778 oryza sativ
P06797 mus musculu
P07154 rattus norv

ALIGNMENTS

[illegible][illegible]

	Query Match	3-44	Scores 110, 57	DB 1	Length 3, 27
Post Local Similarity	41.66				
Matches	462	Conservative	481	Mismatches	491
				Index	3
				Gaps	27

HERBELL 4			
ANAN	ANAN		
115	ANAN	STANDARD	PR1
ANAN	ANAN		54% AA
116	ANAN		
117	ANAN		
118	ANAN		
119	ANAN		
120	ANAN		
121	ANAN		
122	ANAN		
123	ANAN		
124	ANAN		
125	ANAN		
126	ANAN		
127	ANAN		
128	ANAN		
129	ANAN		
130	ANAN		
131	ANAN		
132	ANAN		
133	ANAN		
134	ANAN		
135	ANAN		
136	ANAN		
137	ANAN		
138	ANAN		
139	ANAN		
140	ANAN		
141	ANAN		
142	ANAN		
143	ANAN		
144	ANAN		
145	ANAN		
146	ANAN		
147	ANAN		
148	ANAN		
149	ANAN		
150	ANAN		
151	ANAN		
152	ANAN		
153	ANAN		
154	ANAN		
155	ANAN		
156	ANAN		
157	ANAN		
158	ANAN		
159	ANAN		
160	ANAN		
161	ANAN		
162	ANAN		
163	ANAN		
164	ANAN		
165	ANAN		
166	ANAN		
167	ANAN		
168	ANAN		
169	ANAN		
170	ANAN		
171	ANAN		
172	ANAN		
173	ANAN		
174	ANAN		
175	ANAN		
176	ANAN		
177	ANAN		
178	ANAN		
179	ANAN		
180	ANAN		
181	ANAN		
182	ANAN		
183	ANAN		
184	ANAN		
185	ANAN		
186	ANAN		
187	ANAN		
188	ANAN		
189	ANAN		
190	ANAN		
191	ANAN		
192	ANAN		
193	ANAN		
194	ANAN		
195	ANAN		
196	ANAN		
197	ANAN		
198	ANAN		
199	ANAN		
200	ANAN		

OX N-BL LACTID 44153
PP 11
PP SEQUENCE FROM N.A.
PC STAPIN V. SMITH (CAYMANE) LINDSEY STONE
PC ROBERTSON C.E. Goodson, L.W.
KA "Isolation and expression of monomeric form of papain."
R1 submitted (R.N. 1997) to the 1998/2000 annual meetings.
R2
R3
R4
R5
R6
R7
R8
R9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
R20
R21
R22
R23
R24
R25
R26
R27
R28
R29
R30
R31
R32
R33
R34
R35
R36
R37
R38
R39
R40
R41
R42
R43
R44
R45
R46
R47
R48
R49
R50
R51
R52
R53
R54
R55
R56
R57
R58
R59
R60
R61
R62
R63
R64
R65
R66
R67
R68
R69
R70
R71
R72
R73
R74
R75
R76
R77
R78
R79
R80
R81
R82
R83
R84
R85
R86
R87
R88
R89
R90
R91
R92
R93
R94
R95
R96
R97
R98
R99
R100
R101
R102
R103
R104
R105
R106
R107
R108
R109
R110
R111
R112
R113
R114
R115
R116
R117
R118
R119
R120
R121
R122
R123
R124
R125
R126
R127
R128
R129
R130
R131
R132
R133
R134
R135
R136
R137
R138
R139
R140
R141
R142
R143
R144
R145
R146
R147
R148
R149
R150
R151
R152
R153
R154
R155
R156
R157
R158
R159
R160
R161
R162
R163
R164
R165
R166
R167
R168
R169
R170
R171
R172
R173
R174
R175
R176
R177
R178
R179
R180
R181
R182
R183
R184
R185
R186
R187
R188
R189
R190
R191
R192
R193
R194
R195
R196
R197
R198
R199
R200
R201
R202
R203
R204
R205
R206
R207
R208
R209
R210
R211
R212
R213
R214
R215
R216
R217
R218
R219
R220
R221
R222
R223
R224
R225
R226
R227
R228
R229
R230
R231
R232
R233
R234
R235
R236
R237
R238
R239
R240
R241
R242
R243
R244
R245
R246
R247
R248
R249
R250
R251
R252
R253
R254
R255
R256
R257
R258
R259
R260
R261
R262
R263
R264
R265
R266
R267
R268
R269
R270
R271
R272
R273
R274
R275
R276
R277
R278
R279
R280
R281
R282
R283
R284
R285
R286
R287
R288
R289
R290
R291
R292
R293
R294
R295
R296
R297
R298
R299
R300
R301
R302
R303
R304
R305
R306
R307
R308
R309
R310
R311
R312
R313
R314
R315
R316
R317
R318
R319
R320
R321
R322
R323
R324
R325
R326
R327
R328
R329
R330
R331
R332
R333
R334
R335
R336
R337
R338
R339
R340
R341
R342
R343
R344
R345
R346
R347
R348
R349
R350
R351
R352
R353
R354
R355
R356
R357
R358
R359
R360
R361
R362
R363
R364
R365
R366
R367
R368
R369
R370
R371
R372
R373
R374
R375
R376
R377
R378
R379
R380
R381
R382
R383
R384
R385
R386
R387
R388
R389
R390
R391
R392
R393
R394
R395
R396
R397
R398
R399
R400
R401
R402
R403
R404
R405
R406
R407
R408
R409
R410
R411
R412
R413
R414
R415
R416
R417
R418
R419
R420
R421
R422
R423
R424
R425
R426
R427
R428
R429
R430
R431
R432
R433
R434
R435
R436
R437
R438
R439
R440
R441
R442
R443
R444
R445
R446
R447
R448
R449
R450
R451
R452
R453
R454
R455
R456
R457
R458
R459
R460
R461
R462
R463
R464
R465
R466
R467
R468
R469
R470
R471
R472
R473
R474
R475
R476
R477
R478
R479
R480
R481
R482
R483
R484
R485
R486
R487
R488
R489
R490
R491
R492
R493
R494
R495
R496
R497
R498
R499
R500
R501
R502
R503
R504
R505
R506
R507
R508
R509
R510
R511
R512
R513
R514
R515
R516
R517
R518
R519
R520
R521
R522
R523
R524
R525
R526
R527
R528
R529
R530
R531
R532
R533
R534
R535
R536
R537
R538
R539
R540
R541
R542
R543
R544
R545
R546
R547
R548
R549
R550
R551
R552
R553
R554
R555
R556
R557
R558
R559
R560
R561
R562
R563
R564
R565
R566
R567
R568
R569
R570
R571
R572
R573
R574
R575
R576
R577
R578
R579
R580
R581
R582
R583
R584
R585
R586
R587
R588
R589
R590
R591
R592
R593
R594
R595
R596
R597
R598
R599
R600
R601
R602
R603
R604
R605
R606
R607
R608
R609
R610
R611
R612
R613
R614
R615
R616
R617
R618
R619
R620
R621
R622
R623
R624
R625
R626
R627
R628
R629
R630
R631
R632
R633
R634
R635
R636
R637
R638
R639
R640
R641
R642
R643
R644
R645
R646
R647
R648
R649
R650
R651
R652
R653
R654
R655
R656
R657
R658
R659
R660
R661
R662
R663
R664
R665
R666
R667
R668
R669
R670
R671
R672
R673
R674
R675
R676
R677
R678
R679
R680
R681
R682

[illegible]

FI	66-DEP	25	127
FI	CHAIN	128	349
FT	A ¹ -SITE	147	147
FT	A ¹ -SITE	276	276
FT	B ¹ -SITE	144	164
FT	SOL-FID	176	217
FT	SOL-FID	272	374
FT	CONECTED	291	291
FT	CHAINS	324	324
FT	STAINING	349	AA:
FT	STAINING	349	AB MB: FALCEROCERILADIP-CHEO 4:

[illegible]

[illegible][illegible]


```

Query Match      15.4% Score 137; DB 1; Length 356;
Best local Similarity 30.6%; Prod. No. 0.00082;
Matches 37; Conservative 29; Mismatches 35; Indels 20; Gaps 4;

QY 34 KIKNIITATIAIFIVLVIVSVLYITNTSDKIDDEVPDGYVDPATREYKES---PEEPK 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15 KLOPHIVCTM---NIGIVTALCSAASAAD-----FGAAYNIQRAPDYVESFV 60

QY 51 KKYHUVYSSMEENQREIVKQNFPIKTN---SQCFSVLPMEFCHISFFPPMAPPT 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 ENKKNYTSOWEKNKRYSFKNLHEINAKNUNAIIDPATIYKINKEISLKSSELIAPET 120

QY 148 G 148
   I
DB 121 G 121

RESULT 4
ACTIN.ACTCH STANDARD; PRT; 480 AA.
AC P00785;
DI 21-JUL-1986 (Rel. 01, Created)
DI 01-JAN-1990 (Rel. 13, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Actinidin precursor (EC 3.4.22.14) (Actinidin) (Allergen Act c 1).
OS Actinidia chinensis (Kiwi) (Yangtiao).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Asteridales; Ericales; Actinidiales; Actinidia.
OX NCBI_Taxid:4625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, DELICIOSA;
RX MEDLINE=90045955; PubMed=2813065;
EA P. L. KESSLER, F. E. J. J. VAN DER BEEK, 5.7.77.
RT "Nucleotide sequence of actinidin, a kiwi fruit protease."
RL Nucleic Acids Res. 17:833-836 (1989)
RN [2]
RP SEQUENCE OF 70-371 FROM N.A.
RA Praekelt, G. M., McKee, K. A., Smith, B. J.
RT "Molecular analysis of actinidin, the cysteine proteinase of Actinidia chinensis."
RL Plant Mol. Biol. 10:193-202 (1988).
RN [3]
RP SEQUENCE OF 127-246.
RX MEDLINE=90062777; PubMed=697390.
RA Carne, A., Moore, G. H. J.
RT "The amino acid sequence of the tryptic peptides from actinidin, a proteolytic enzyme from the fruit of Actinidia chinensis."
RL Biochem. J. 173:773-83 (1978).
RN [4]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=91346716; PubMed=2102886;
RA Keeling, J., Maxwell, P., Gardner, R. C. J.
RT "Nucleotide sequence of the promoter region from kiwifruit actinidin genes."
RL Plant Mol. Biol. 15:787-788 (1990).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS), AND REVISIONS.
RX MEDLINE=R1072298; PubMed=7903158;
RA BAYER, F. N.
RT "Structure of actinidin, after refinement at 1.7-A resolution."
RL J. Mol. Biol. 141:441-484 (1980).
CC -1- CATALYTIC ACTIVITY: Specificity close to that of papain.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE PAPAINE FAMILY OF THIOL PROTEASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ambassy/)

```

```

or send an email to license@sib-sib.ch.
-----
CC EMBL; X16466; CAA34486.1; -
CC EMBL; X13013; CAA31435.1; -
DR EMBL; X13139; CAA31529.1; -
DR EMBL; X57551; CAA40778.1; -
DR PIR; A00975; TAGB.
DR PIR; S06587; S06587.
DR PSP; DACT; 15 JAN 86.
DR MEROPS; C01.007; -.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Thiolprot_act_site.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAINE.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR Hydrolase; Thiol; Protease; Zymogen; Signal; Allergen; 3D-structure.
KW SIGNAL.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 126 ACTIVATION PEPTIDE.
FT CHAIN 127 380 ACTINIDAIN.
FT ACT_SITE 151 151
FT ACT_SITE 288 288
FT ACT_SITE 308 308
FT DISULFID 148 191
FT DISULFID 182 224
FT DISULFID 282 342
FT CONFLICT 90 95 E -> G (IN REF. 2).
FT CONFLICT 108 108 S -> G (IN REF. 2).
FT CONFLICT 123 124 PG -> VS (IN REF. 2).
FT CONFLICT 167 170 VTGV -> TSGS (IN REF. 3).
FT CONFLICT 181 181 D -> G (IN REF. 2).
FT CONFLICT 184 184 F -> G (IN REF. 2).
FT CONFLICT 192 192 N -> D (IN REF. 3).
FT CONFLICT 206 206 N -> D (IN REF. 3).
FT CONFLICT 212 212 E -> G (IN REF. 2).
FT CONFLICT 223 227 EGNLD -> LQDVA (IN REF. 3).
FT CONFLICT 230 231 NE -> DQ (IN REF. 3).
FT CONFLICT 240 240 E -> G (IN REF. 2).
FT CONFLICT 272 274 HYS -> OVA (IN REF. 3).
FT CONFLICT 282 285 I -> V (IN REF. 3).
FT CONFLICT 290 291 VT -> IV (IN REF. 3).
FT CONFLICT 301 301 I -> V (IN REF. 3).
FT CONFLICT 307 307 K -> E (IN REF. 2).
FT CONFLICT 349 349 H -> Y (IN REF. 2).
FT CONFLICT 360 360 P -> S (IN REF. 2).
FT CONFLICT 371 380 GVDDGQRYSA -> E (IN REF. 2).
FT STRAND 141 192
FT HELIX 133 136
FT TURN 137 137
FT STRAND 144 144
FT TURN 146 147
FT HELIX 151 168
FT STRAND 174 174
FT HELIX 176 182
FT STRAND 185 185
FT TURN 186 187
FT STRAND 188 188
FT HELIX 190 192
FT TURN 196 206
FT TURN 207 207
FT STRAND 209 209
FT STRAND 211 211
FT TURN 212 214
FT HELIX 226 230
FT STRAND 234 234
FT STRAND 238 241
FT TURN 244 245
FT HELIX 247 256
FT STRAND 259 263
FT HELIX 268 272
FT STRAND 277 278
FT STRAND 294 298

```


11 100N 499 607
F1 SIPIANE 601 607
F1 SIPIANE 612 618
F1 100N 615 621
F1 SIPIANE 616 622
F1 100N 617 623
F1 SIPIANE 619 625
F1 100N 626 632
F1 HELIX 631 637
F1 100N 634 640
F1 SIPIANE 639 645
SQ SEQUENCE 642 648 649 655 661 667 673 679 685 691 697 703 709 715 721 727 733 739 745 751 757 763 769 775 781 787 793 799 805 811 817 823 829 835 841 847 853 859 865 871 877 883 889 895 901 907 913 919 925 931 937 943 949 955 961 967 973 979 985 991 997 1003 1009 1015 1021 1027 1033 1039 1045 1051 1057 1063 1069 1075 1081 1087 1093 1099 1105 1111 1117 1123 1129 1135 1141 1147 1153 1159 1165 1171 1177 1183 1189 1195 1201 1207 1213 1219 1225 1231 1237 1243 1249 1255 1261 1267 1273 1279 1285 1291 1297 1303 1309 1315 1321 1327 1333 1339 1345 1351 1357 1363 1369 1375 1381 1387 1393 1399 1405 1411 1417 1423 1429 1435 1441 1447 1453 1459 1465 1471 1477 1483 1489 1495 1501 1507 1513 1519 1525 1531 1537 1543 1549 1555 1561 1567 1573 1579 1585 1591 1597 1603 1609 1615 1621 1627 1633 1639 1645 1651 1657 1663 1669 1675 1681 1687 1693 1699 1705 1711 1717 1723 1729 1735 1741 1747 1753 1759 1765 1771 1777 1783 1789 1795 1801 1807 1813 1819 1825 1831 1837 1843 1849 1855 1861 1867 1873 1879 1885 1891 1897 1903 1909 1915 1921 1927 1933 1939 1945 1951 1957 1963 1969 1975 1981 1987 1993 1999 2005 2011 2017 2023 2029 2035 2041 2047 2053 2059 2065 2071 2077 2083 2089 2095 2101 2107 2113 2119 2125 2131 2137 2143 2149 2155 2161 2167 2173 2179 2185 2191 2197 2203 2209 2215 2221 2227 2233 2239 2245 2251 2257 2263 2269 2275 2281 2287 2293 2299 2305 2311 2317 2323 2329 2335 2341 2347 2353 2359 2365 2371 2377 2383 2389 2395 2401 2407 2413 2419 2425 2431 2437 2443 2449 2455 2461 2467 2473 2479 2485 2491 2497 2503 2509 2515 2521 2527 2533 2539 2545 2551 2557 2563 2569 2575 2581 2587 2593 2599 2605 2611 2617 2623 2629 2635 2641 2647 2653 2659 2665 2671 2677 2683 2689 2695 2701 2707 2713 2719 2725 2731 2737 2743 2749 2755 2761 2767 2773 2779 2785 2791 2797 2803 2809 2815 2821 2827 2833 2839 2845 2851 2857 2863 2869 2875 2881 2887 2893 2899 2905 2911 2917 2923 2929 2935 2941 2947 2953 2959 2965 2971 2977 2983 2989 2995 3001 3007 3013 3019 3025 3031 3037 3043 3049 3055 3061 3067 3073 3079 3085 3091 3097 3103 3109 3115 3121 3127 3133 3139 3145 3151 3157 3163 3169 3175 3181 3187 3193 3199 3205 3211 3217 3223 3229 3235 3241 3247 3253 3259 3265 3271 3277 3283 3289 3295 3301 3307 3313 3319 3325 3331 3337 3343 3349 3355 3361 3367 3373 3379 3385 3391 3397 3403 3409 3415 3421 3427 3433 3439 3445 3451 3457 3463 3469 3475 3481 3487 3493 3499 3505 3511 3517 3523 3529 3535 3541 3547 3553 3559 3565 3571 3577 3583 3589 3595 3601 3607 3613 3619 3625 3631 3637 3643 3649 3655 3661 3667 3673 3679 3685 3691 3697 3703 3709 3715 3721 3727 3733 3739 3745 3751 3757 3763 3769 3775 3781 3787 3793 3799 3805 3811 3817 3823 3829 3835 3841 3847 3853 3859 3865 3871 3877 3883 3889 3895 3901 3907 3913 3919 3925 3931 3937 3943 3949 3955 3961 3967 3973 3979 3985 3991 3997 4003 4009 4015 4021 4027 4033 4039 4045 4051 4057 4063 4069 4075 4081 4087 4093 4099 4105 4111 4117 4123 4129 4135 4141 4147 4153 4159 4165 4171 4177 4183 4189 4195 4201 4207 4213 4219 4225 4231 4237 4243 4249 4255 4261 4267 4273 4279 4285 4291 4297 4303 4309 4315 4321 4327 4333 4339 4345 4351 4357 4363 4369 4375 4381 4387 4393 4399 4405 4411 4417 4423 4429 4435 4441 4447 4453 4459 4465 4471 4477 4483 4489 4495 4501 4507 4513 4519 4525 4531 4537 4543 4549 4555 4561 4567 4573 4579 4585 4591 4597 4603 4609 4615 4621 4627 4633 4639 4645 4651 4657 4663 4669 4675 4681 4687 4693 4699 4705 4711 4717 4723 4729 4735 4741 4747 4753 4759 4765 4771 4777 4783 4789 4795 4801 4807 4813 4819 4825 4831 4837 4843 4849 4855 4861 4867 4873 4879 4885 4891 4897 4903 4909 4915 4921 4927 4933 4939 4945 4951 4957 4963 4969 4975 4981 4987 4993 4999 5005 5011 5017 5023 5029 5035 5041 5047 5053 5059 5065 5071 5077 5083 5089 5095 5101 5107 5113 5119 5125 5131 5137 5143 5149 5155 5161 5167 5173 5179 5185 5191 5197 5203 5209 5215 5221 5227 5233 5239 5245 5251 5257 5263 5269 5275 5281 5287 5293 5299 5305 5311 5317 5323 5329 5335 5341 5347 5353 5359 5365 5371 5377 5383 5389 5395 5401 5407 5413 5419 5425 5431 5437 5443 5449 5455 5461 5467 5473 5479 5485 5491 5497 5503 5509 5515 5521 5527 5533 5539 5545 5551 5557 5563 5569 5575 5581 5587 5593 5599 5605 5611 5617 5623 5629 5635 5641 5647 5653 5659 5665 5671 5677 5683 5689 5695 5701 5707 5713 5719 5725 5731 5737 5743 5749 5755 5761 5767 5773 5779 5785 5791 5797 5803 5809 5815 5821 5827 5833 5839 5845 5851 5857 5863 5869 5875 5881 5887 5893 5899 5905 5911 5917 5923 5929 5935 5941 5947 5953 5959 5965 5971 5977 5983 5989 5995 6001 6007 6013 6019 6025 6031 6037 6043 6049 6055 6061 6067 6073 6079 6085 6091 6097 6103 6109 6115 6121 6127 6133 6139 6145 6151 6157 6163 6169 6175 6181 6187 6193 6199 6205 6211 6217 6223 6229 6235 6241 6247 6253 6259 6265 6271 6277 6283 6289 6295 6301 6307 6313 6319 6325 6331 6337 6343 6349 6355 6361 6367 6373 6379 6385 6391 6397 6403 6409 6415 6421 6427 6433 6439 6445 6451 6457 6463 6469 6475 6481 6487 6493 6499 6505 6511 6517 6523 6529 6535 6541 6547 6553 6559 6565 6571 6577 6583 6589 6595 6601 6607 6613 6619 6625 6631 6637 6643 6649 6655 6661 6667 6673 6679 6685 6691 6697 6703 6709 6715 6721 6727 6733 6739 6745 6751 6757 6763 6769 6775 6781 6787 6793 6799 6805 6811 6817 6823 6829 6835 6841 6847 6853 6859 6865 6871 6877 6883 6889 6895 6901 6907 6913 6919 6925 6931 6937 6943 6949 6955 6961 6967 6973 6979 6985 6991 6997 7003 7009 7015 7021 7027 7033 7039 7045 7051 7057 7063 7069 7075 7081 7087 7093 7099 7105 7111 7117 7123 7129 7135 7141 7147 7153 7159 7165 7171 7177 7183 7189 7195 7201 7207 7213 7219 7225 7231 7237 7243 7249 7255 7261 7267 7273 7279 7285 7291 7297 7303 7309 7315 7321 7327 7333 7339 7345 7351 7357 7363 7369 7375 7381 7387 7393 7399 7405 7411 7417 7423 7429 7435 7441 7447 7453 7459 7465 7471 7477 7483 7489 7495 7501 7507 7513 7519 7525 7531 7537 7543 7549 7555 7561 7567 7573 7579 7585 7591 7597 7603 7609 7615 7621 7627 7633 7639 7645 7651 7657 7663 7669 7675 7681 7687 7693 7699 7705 7711 7717 7723 7729 7735 7741 7747 7753 7759 7765 7771 7777 7783 7789 7795 7801 7807 7813 7819 7825 7831 7837 7843 7849 7855 7861 7867 7873 7879 7885 7891 7897 7903 7909 7915 7921 7927 7933 7939 7945 7951 7957 7963 7969 7975 7981 7987 7993 7999 8005 8011 8017 8023 8029 8035 8041 8047 8053 8059 8065 8071 8077 8083 8089 8095 8101 8107 8113 8119 8125 8131 8137 8143 8149 8155 8161 8167 8173 8179 8185 8191 8197 8203 8209 8215 8221 8227 8233 8239 8245 8251 8257 8263 8269 8275 8281 8287 8293 8299 8305 8311 8317 8323 8329 8335 8341 8347 8353 8359 8365 8371 8377 8383 8389 8395 8401 8407 8413 8419 8425 8431 8437 8443 8449 8455 8461 8467 8473 8479 8485 8491 8497 8503 8509 8515 8521 8527 8533 8539 8545 8551 8557 8563 8569 8575 8581 8587 8593 8599 8605 8611 8617 8623 8629 8635 8641 8647 8653 8659 8665 8671 8677 8683 8689 8695 8701 8707 8713 8719 8725 8731 8737 8743 8749 8755 8761 8767 8773 8779 8785 8791 8797 8803 8809 8815 8821 8827 8833 8839 8845 8851 8857 8863 8869 8875 8881 8887 8893 8899 8905 8911 8917 8923 8929 8935 8941 8947 8953 8959 8965 8971 8977 8983 8989 8995 9001 9007 9013 9019 9025 9031 9037 9043 9049 9055 9061 9067 9073 9079 9085 9091 9097 9103 9109 9115 9121 9127 9133 9139 9145 9151 9157 9163 9169 9175 9181 9187 9193 9199 9205 9211 9217 9223 9229 9235 9241 9247 9253 9259 9265 9271 9277 9283 9289 9295 9301 9307 9313 9319 9325 9331 9337 9343 9349 9355 9361 9367 9373 9379 9385 9391 9397 9403 9409 9415 9421 9427 9433 9439 9445 9451 9457 9463 9469 9475 9481 9487 9493 9499 9505 9511 9517 9523 9529 9535 9541 9547 9553 9559 9565 9571 9577 9583 9589 9595 9601 9607 9613 9619 9625 9631 9637 9643 9649 9655 9661 9667 9673 9679 9685 9691 9697 9703 9709 9715 9721 9727 9733 9739 9745 9751 9757 9763 9769 9775 9781 9787 9793 9799 9805 9811 9817 9823 9829 9835 9841 9847 9853 9859 9865 9871 9877 9883 9889 9895 9901 9907 9913 9919 9925 9931 9937 9943 9949 9955 9961 9967 9973 9979 9985 9991 9997 10003 10009 10015 10021 10027 10033 10039 10045 10051 10057 10063 10069 10075 10081 10087 10093 10099 10105 10111 10117 10123 10129 10135 10141 10147 10153 10159 10165 10171 10177 10183 10189 10195 10201 10207 10213 10219 10225 10231 10237 10243 10249 10255 10261 10267 10273 10279 10285 10291 10297 10303 10309 10315 10321 10327 10333 10339 10345 10351 10357 10363 10369 10375 10381 10387 10393 10399 10405 10411 10417 10423 10429 10435 10441 10447 10453 10459 10465 10471 10477 10483 10489 10495 10501 10507 10513 10519 10525 10531 10537 10543 10549 10555 10561 10567 10573 10579 10585 10591 10597 10603 10609 10615 10621 10627 10633 10639 10645 10651 10657 10663 10669 10675 10681 10687 10693 10699 10705 10711 10717 10723 10729 10735 10741 10747 10753 10759 10765 10771 10777 10783 10789 10795 10801 10807 10813 10819 10825 10831 10837 10843 10849 10855 10861 10867 10873 10879 10885 10891 10897 10903 10909 10915 10921 10927 10933 10939 10945 10951 10957 10963 10969 10975 10981 10987 10993 10999 11005 11011 11017 11023 11029 11035 11041 11047 11053 11059 11065 11071 11077 11083 11089 11095 11101 11107 11113 11119 11125 11131 11137 11143 11149 11155 11161 11167 11173 11179 11185 11191 11197 11203 11209 11215 11221 11227 11233 11239 11245 11251 11257 11263 11269 11275 11281 11287 11293 11299 11305 11311 11317 11323 11329 11335 11341 11347 11353 11359 11365 11371 11377 11383 11389 11395 11401 11407 11413 11419 11425 11431 11437 11443 11449 11455 11461 11467 11473 11479 11485 11491 11497 11503 11509 11515 11521 11527 11533 11539 11545 11551 11557 11563 11569 11575 11581 11587 11593 11599 11605 11611 11617 11623 11629 11635 11641 11647 11653 11659 11665 11671 11677 11683 11689 11695 11701 11707 11713 11719 11725 11731 11737 11743 11749 11755 11761 11767 11773 11779 11785 11791 11797 11803 11809 11815 11821 11827 11833 11839 11845 11851 11857 11863 11869 11875 11881 11887 11893 11899 11905 11911 11917 11923 11929 11935 11941 11947 11953 11959 11965 11971 11977 11983 11989 11995 12001 12007 12013 12019 12025 12031 12037 12043 12049 12055 12061 12067 12073 12079 12085 12091 12097 12103 12109 12115 12121 12127 12133 12139 12145 12151 12157 12163 12169 12175 12181 12187 12193 12199 12205 12211 12217 12223 12229 12235 12241 12247 12253 12259 12265 12271 12277 12283 12289 12295 12301 12307 12313 12319 12325 12331 12337 12343 12349 12355 12361 12367 12373 12379 12385 12391 12397 12403 12409 12415 12421 12427 12433 12439 12445 12451 12457 12463 12469 12475 12481 12487 12493 12499 12505 12511 12517 12523 12529 12535 12541 12547 12553 12559 12565 12571 12577 12583 12589 12595 12601 12607 12613 12619 12625 12631 12637 12643 12649 12655 12661 12667 12673 12679 12685 12691 12697 12703 12709 12715 12721 12727 12733 12739 12745 12751 12757 12763 12769 12775 12781 12787 12793 12799 12805 12811 12817 12823 12829 12835 12841 12847 12853 12859 12865 12871 12877 12883 12889 12895 12901 12907 12913 12919 12925 12931 12937 12943 12949 12955 12961 12967 12973 12979 12985 12991 12997 13003 13009 13015 13021 13027 13033 13039 13045 13051 13057 13063 13069 13075 13081 13087 13093 13099 13105 13111 13117 13123 13129 13135 13141 13147 13153 13159 13165 13171 13177 13183 13189 13195 13201 13207 13213 13219 13225 13231 13237 13243 13249 13255 13261 13267 13273 13279 13285 13291 13297 13303 13309 13315 13321 13327 13333 13339 13345 13351 13357 13363 13369 13375 13381 13387 13393 13399 13405 13411 13417 13423 13429 13435 13441 13447 13453 13459 13465 13471 13477 13483 13489 13495 13501 13507 13513 13519 13525 13531 13537 13543 13549 13555 13561 13567 13573 13579 13585 13591 13597 13603 13609 13615 13621 13627 13633 13639 13645 13651 13657 13663 13669 13675 13681 13687 13693 13699 13705 13711 13717 13723 13729 13735 13741 13747 13753 13759 13765 13771 13777 13783 13789 13795 13801 13807 13813 13819 13825 13831 13837 13843 13849 13855 13861 13867 13873 13879 13885 13891 13897 13903 13909 13915 13921 13927 13933 13939 13945 13951 13957 13963 13969 13975 13981 13987 13993 13999 14005 14011 14017 14023 14029 14035 14041 14047 14053 14059 14065 14071 14077 14083 14089 14095 14101 14107 14113 14119 14125 14131 14137 14143 14149 14155 14161 14167 14173 14179 14185 14191 14197 14203 14209 14215 14221 14227 14233 14239 14245 14251 14257 14263 14269 14275 14281 14287 14293 1429

FT DISULFID 167 207 BY SIMILARITY.
 FT DISULFID 242 316 BY SIMILARITY.
 SQ SEQUENCE 324 AA: 26422 MW: 48027.777394 0.9564.

Query Match 15.0% Score 134.5; DB 1; Length 323;
 Best local Similarity 34.0%; Pred No 0 0012;
 Matches 36; Conservative 19; Mismatches 30; Indels 21; Gaps 3;

Q) 43 LIAF:VLVTIVSYLTINFSPTFIVVGVVEFATFVPSFTFFFFVYKVVSSMRF 102
 DB 5 LYLEFVAVKSAAY-----DKIKAPNY-----FEFVHREKNYSSEVE 44
 QY 103 ENQFFVYKNNFIKTNQSGSYVLENNFGDLSKEEFMARF 148
 DB 45 KLEPFKIFQHNELINK-EN-SAYRINKISLSEKELIAYTG 89

RESULT 12
 CYSPT_THFAN
 ID CYSPT_THFAN STANDARD; PRT: 441 AA.
 AC P25781;
 DT 01-MAY-1992 (Ref. 22; Created)
 DI 01-MAY-1992 (Ref. 22; Last sequence update)
 DE 01-NOV-1995 (Ref. 32; Last annotation update)
 DE Cysteine proteinase precursor (NC 3.4.22.1)
 GN TACP.
 OS Theileria annulata
 OC Eukaryota; Alveolata; Apicomplexa; Eimeriidae; Theileridae;
 CC Theileria.
 CX NCBI_TaxID=5874;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92389980; PubMed=1518523;
 RA Baylis H.A., Meeson A., Mottram J.C., Hall R.;
 RT "Characterisation of a gene for a cysteine proteinase from Theileria
 annulata";
 RL Mol. Biochem. Parasitol. 64:105-107(1992)
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 PAPAEN FAMILY OF THIOL PROTEASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announc-7>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M86659; AAA30135.1;
 DR PIR: A45555; A45555;
 DR HSSP: P14080; IVAL;
 DR MEROPS: C01.073; 7;
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprot_act_site.
 DE PRAM: PR0122; Peptidase_C1; 2;
 DR PRINTS: PR00705; PAPAEN.
 DR PROSITE: PS00130; THIOL_PROTEASE_CYS; FALSE_NEG.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1;
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1;
 DR Hydrolase: thiol protease; 2; Proton; Signal; Glycoprotein.
 KW SIGNAL
 FT PROPEP 1 ?
 FT CHAIN 229 411
 FT ACT_SITE 252 252
 FT ACT_SITE 381 381 BY SIMILARITY.
 FT ACT_SITE 403 403 BY SIMILARITY.
 FT ACT_SITE 249 290 BY SIMILARITY.
 FT DISULFID 249 290
 FT CARBOHYD 345 345 N-LINKED (GLNAC...) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLNAC...) (POTENTIAL).
 SQ SEQUENCE 441 AA: 45653 MW: 65674.65521330 0.9564.

Query Match 15.0% Score 134.5; DB 1; Length 441;
 Best local Similarity 27.2%; Pred No 0 0016;
 Matches 42; Conservative 30; Mismatches 38; Indels 30; Gaps 5;

QY 36 YNIIATLIAIFIVIVTVSY-----ITNNTS-----DKTDQVPGVGVVDPATP--- 80
 DB 36 KLLISLPLTFILALSSASILTFFFRSKSITNFKSLAIEHTESHYS--MDPSKRAGF 93
 QY 81 -----EYRKS-----FEFEKKYHKVYSSMEENORFEIYKNNNFKT 119
 DB 94 VEEIVKIQGKITSRAESELMIIEFDAPVFRYKRVHPSFDQVDFELTFRKNTHIVK- 152
 QY 120 TNSGFSYVLENNFGDLSKEEFMARF 146
 DB 153 THRTEPYSLDLNKFSDLSDEEFKALY 179

RESULT 13
 PAPA_CARPA
 ID PAPA_CARPA STANDARD; PRT: 348 AA.
 AC P05954;
 DT 01-NOV-1988 (Ref. 09; Created)
 DI 01-OCT-1996 (Ref. 34; Last sequence update)
 DI 11-JUL-2001 (Ref. 43; Last annotation update)
 DE papaya proteinase IV precursor (EC 3.4.22.25) (pPIV) (papaya peptidase
 B) (Glycyl endopeptidase).
 OS Carica papaya (Papaya).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CX NCBI_TaxID=3649;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Leaf;
 RX MEDLINE=95288290; PubMed=7770454;
 RA Taylor M.A., Bacci F.C., Briggs G.S., Conneron I.P., Cummings N.J.,
 Pratt K.A., Peyrell P., Freedman P.R., Goodenough P.W.;
 RT "Purification and characterization of papaya proteinase IV are
 selective high affinity inhibitors of the mature papaya enzymes";
 RL Protein Eng. 8:59-62(1995).
 RN [2]
 RN SEQUENCE OF 133-348.
 RX MEDLINE=90076470; PubMed=2591528;
 RA Ritonja A., Buttle D.J., Pawlings N.D., Turk V., Barrett A.J.;
 RT "Papaya proteinase IV amino acid sequence";
 RL FEBS Lett. 258:109-112(1989).
 RN [3]
 RP SEQUENCE OF 133-150.
 RX MEDLINE=80088401; PubMed=518921;
 RA Lynn K.R., Yaguchi M.;
 RT "N terminal homology in three cysteinyl proteases from Papaya latex";
 RL Biochim. Biophys. Acta 581:363-364(1979).
 RN [4]
 RP SEQUENCE OF 287-348 FROM N.A.
 RX MEDLINE=87099799; PubMed=3541893;
 RA McKee K.A., Adams S., Matthews J.A., Smith C.J., Smith H.;
 RT "Molecular cloning of two cysteine proteinases from paw-paw (Carica
 papaya)";
 RL Biochem. J. 237:105-110(1986).
 RN [5]
 RP SUBSTRATE SPECIFICITY.
 RX MEDLINE=93127127; PubMed=2404757;
 RA Ritonja A., Buttle D.J., Turk V., Barrett A.J.;
 RT "Selective cleavage of glycosyl bands by papaya proteinase IV";
 RL FEBS Lett. 260:195-197(1990).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=96006178; PubMed=7548082;
 RA Gohara S., Homma A.M., Buttle D.J., Pearl L.H.;
 RT "Crystal structure of glycol endopeptidase from Carica papaya: a
 cysteine endopeptidase of unusual substrate specificity";
 RL Biochemistry 34:13190-13195(1995).
 CC -1- FUNCTION: THIOL PROTEASE WITH A SUBSTRATE SPECIFICITY VERY

CC -1- DEVELOPMENTAL STAGE: THROPHOZOID.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAININ FAMILY OF THIOL PROTEASES.
 CC -1- SIMILARITY: STRONGEST SIMILARITY TO CATHEPSIN L.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1st-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M81341; AAA29578.1; -;
 DR PIR: A45624; A45624.
 DR HSP: P07688; 1QDQ.
 DR MEROPS: C01.077; -;
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprot_act_site.
 DR Pfam: PF00112; peptidase_C1_2.
 DR PRINTS: PR00705; PAININ
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPE 2 332 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 333 569 THIOPHOZOITE CYSTEINE PROTEINASE.
 FT DOMAIN 64 70 POLY-ASN.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 488 488 BY SIMILARITY.
 FT ACT_SITE 533 533 BY SIMILARITY.
 FT DISULFID 354 395 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 569 AA; 66880 MW; E5E762A3637E7060 CRC64;

Query Match 14.88; Score 133; DR 1; Length 569;
 Best Local Similarity 27.18; Pred. No. 9.0028;
 Matches 46; Conservative 34; Mismatches 60; Indels 30; Gaps 7;
 QY 5 RSV - EERGEYS GPVALININ-REFENKELIATLIAIV 49
 DB 120 RYVSHFNSKPFPEYELIILLEYKPTENNE -ENRERSLKELIINKLI 177
 QY 50 L---VVIVSLYITNISKD-----IDFVDEYVDVATF-EYKSEPEFKKVKVYS 98
 DB 178 LKGEELKREKVNCHVDFPA-ETSMFFMPYFETINR-FVASELIFKEEINKVKR 237
 QY 99 SMREPNOREFHYKONNPIKTN--SQGFSYVLENNEDLSKEPFMAPF 146
 DB 238 NIDQMKKEEIPKINVISIKNINKLNKMYKKVKVQFSYSEELKEYF 287

Search completed: October 9, 2002, 09:01:06
 Job time: 2163 sec

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in financial matters. The text outlines various methods for organizing and storing records, including digital databases and physical filing systems. It also mentions the need for regular audits and reviews to ensure the integrity of the data.

2. The second part of the document focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides guidelines for effective communication, such as using appropriate language, listening actively, and providing feedback. It also discusses the benefits of open communication, including improved collaboration and decision-making.

3. The third part of the document addresses the issue of risk management. It defines risk as the potential for loss or damage and explains how to identify, assess, and mitigate risks. The text provides a framework for risk management, including the identification of risks, the assessment of their likelihood and impact, and the implementation of control measures. It also discusses the importance of monitoring and reviewing risks over time.

4. The fourth part of the document discusses the importance of training and development. It emphasizes that ongoing training and development are essential for maintaining a skilled and motivated workforce. The text outlines various training and development methods, including classroom-based training, on-the-job training, and self-directed learning. It also discusses the importance of setting learning objectives and evaluating the effectiveness of training programs.

5. The fifth part of the document discusses the importance of innovation and creativity. It emphasizes that innovation and creativity are essential for staying competitive in a rapidly changing market. The text provides guidelines for fostering innovation and creativity, such as encouraging open-mindedness, providing resources for experimentation, and rewarding creative ideas. It also discusses the importance of protecting intellectual property and managing innovation projects.

6. The sixth part of the document discusses the importance of sustainability. It defines sustainability as the ability to meet the needs of the present without compromising the ability of future generations to meet their own needs. The text outlines various sustainability initiatives, including energy conservation, waste reduction, and social responsibility. It also discusses the importance of measuring and reporting on sustainability performance.

7. The seventh part of the document discusses the importance of ethics and integrity. It emphasizes that ethics and integrity are essential for building trust and credibility. The text provides guidelines for ethical behavior, such as being honest, transparent, and respectful. It also discusses the importance of establishing a code of ethics and providing training on ethical issues.

8. The eighth part of the document discusses the importance of customer service. It emphasizes that excellent customer service is essential for retaining customers and attracting new ones. The text provides guidelines for providing excellent customer service, such as listening to customer needs, responding promptly, and resolving problems effectively. It also discusses the importance of training customer service staff and monitoring customer satisfaction.

9. The ninth part of the document discusses the importance of financial management. It emphasizes that sound financial management is essential for the long-term success of an organization. The text outlines various financial management practices, including budgeting, forecasting, and financial reporting. It also discusses the importance of maintaining accurate financial records and seeking professional advice when needed.

10. The tenth part of the document discusses the importance of strategic planning. It emphasizes that strategic planning is essential for setting clear goals and determining the best way to achieve them. The text provides guidelines for developing a strategic plan, such as conducting a SWOT analysis, setting SMART goals, and implementing the plan. It also discusses the importance of reviewing and updating the strategic plan regularly.

GenCore version 4.5
Copyright (c) 1993 - 2000 Computer Ltd.

QM protein - protein search, using sw model

Run on: October 9, 2002, 08:59:16 : Search time 434.73 seconds
(without alignments)
74,986 Million cell updates/sec

Title: US-09-598-062-5

Perfect score: 896

Sequence: 1 MLCGNVVEHQYISGPVIA.....DERVSRV:ASKEEEFFV 175

Scoring table:

Gap 10.0 / Gapext 0.5

Searched: 562222 seqs, 17299429 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPINEMBL_v1:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_orquaele:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteria:
17: sp_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	896	100.0	401	5 Q9VIT4	Q9VIT4 cryptospori
2	195	20.6	374	5 Q9VIT4	Q9VIT4 cryptospori
3	173	19.4	364	10 Q24321	Q24321 phaseolus v
4	170	19.0	368	10 Q41696	Q41696 vicia sativ
5	169	18.9	364	10 Q41110	Q41110 phaseolus v
6	166.5	18.6	365	10 Q940H7	Q940H7 oryza sativ
7	163	18.2	357	10 Q42673	Q42673 carica papa
8	158.5	17.7	464	10 Q41064	Q41064 pisum sativ
9	158	17.6	357	10 Q81085	Q81085 ananas como
10	156	17.4	315	10 Q91M66	Q91M66 arabidopsis
11	156	17.4	356	10 Q95YQ2	Q95YQ2 arabidopsis
12	155.5	17.4	224	10 Q94801	Q94801 ananas como
13	155.5	17.4	352	10 Q9SM11	Q9SM11 carica papa
14	155.5	17.4	361	10 Q9SM10	Q9SM10 carica papa
15	155	17.3	341	10 Q9SC15	Q9SC15 arabidopsis
16	155	17.3	352	10 Q24641	Q24641 ananas como

17	154.5	17.2	356	10	Q81084	Q81084 ananas como
18	154.5	17.2	356	10	Q23799	Q23799 ananas como
19	154.5	17.2	358	10	Q91L83	Q91L83 arabidopsis
20	154	17.2	358	10	Q94HK7	Q94HK7 oryza sativ
21	151.5	16.9	343	10	Q9PS02	Q9PS02 astragalus
22	150	16.7	340	10	Q23800	Q23800 ananas como
23	150	16.7	377	10	Q9M3E7	Q9M3E7 arabidopsis
24	150	16.7	455	13	Q24323	Q24323 phaseolus v
25	149.5	16.7	351	10	Q23791	Q23791 ananas como
26	148.5	16.6	179	10	Q945E4	Q945E4 carica cand
27	147	16.4	325	10	Q39451	Q39451 cicer ariet
28	146.5	16.4	359	10	Q40877	Q40877 petunia hyb
29	145	16.2	361	10	Q94VV8	Q94VV8 oryza sativ
30	142.5	15.9	374	10	Q41522	Q41522 triticum ae
31	142	15.8	342	10	Q41721	Q41721 sinia eleg
32	142	15.8	355	10	Q65493	Q65493 arabidopsis
33	142	15.8	367	10	Q40993	Q40993 pisum sativ
34	140.5	15.7	348	10	Q9ZQH7	Q9ZQH7 arabidopsis
35	140	15.6	326	5	Q27760	Q27760 penaeus van
36	140	15.6	328	5	Q27759	Q27759 penaeus van
37	140	15.6	340	10	Q38706	Q38706 alnus glut
38	140	15.6	346	10	Q480H3	Q480H3 brassica na
39	140	15.6	360	10	Q41P12	Q41P12 nicotiana t
40	140	15.6	460	5	Q94504	Q94504 dictyosteli
41	138	15.4	358	10	Q60002	Q60002 prunus arme
42	137.5	15.3	374	10	Q24137	Q24137 nicotiana t
43	137.5	15.3	454	10	Q49922	Q49922 pseudotsuga
44	137	15.3	380	10	Q43467	Q43467 actinidia c
45	136.5	15.2	350	5	Q25888	Q25888 foxocara ca

ALIGNMENTS

RESULT 1

Q9VIT4 PRELIMINARY: PRT; 401 AA.
ID Q9VIT4
AC Q9VIT4;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CRYPTOSPORIDIUM PARVUM.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN NINC;
RA Petersen G., Huang J.-X.;
ET "Characterization and expression of the gene for cryptopain, a
RT Cryptosporidium parvum cysteine proteinase homologue.";
R1 Submitted (SEP 1998) to the EMBL/GenBank/DDBP databases.
DR EMBL: AF091366; AAD42940.1;
DR HSSP: P07711; ICJL.
DR MEROPS: C01.0PA; -;
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000169; Thiolprot_act_site.
DE Pfam: PF02112; Peptidase_C1, 1.
DE PRINTS: PR00705; PAPAIN.
DE PROSITE: PS06640; THIOI_PROTEASE_ASN; 1.
DE PROSITE: PS00149; THIOI_PROTEASE_CYS; 1.
KW Hydrolase; Signal; Thiol protease.
FT SIGNAL 1 175 POTENTIAL.
FT CHAIN 176 401 CRYPTOPAIN.
SQ SEQUENCE 401 AA: 45457 MW: 55440Da H2A61R70 CXC64;

Query Match: 100.0%; Score 896; DB 5; Length 401;
Best Local Similarity: 100.0%; Prog. No. 6.5e-61;
Matches: 175; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 1 MLCGNVVEHQYISGPVIALINGTNGQEPNKKLNIIATLIAIFIVLVTVTSLYTN 60


```

FT DISULFID      146      188           BY SIMILARITY.
FT DISULFID     180      221           BY SIMILARITY.
FT DISULFID     279      331           BY SIMILARITY.
FT CAGEHEAD      26       26           N-LINKED (GLCNAC... ) (POTENTIAL).
FT CAGEHEAD      74       74           N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE      64 AA, 4-5474 MW, 27416 Da [uniprot] 18747 18764;

Query Match              18.9%; Score 169; DB 10; Length 364;
Best Local Similarity    26.6%; Pred. No. 2.1e-05;
Matches 34, Conservative 36, Mismatches 40, Indels 18, Gaps 2;

QY   35 LKNIIITLAIFLVIVWISLYIINNTSUKIDDFVGDIYPATPKSFEEFKKKYH 94
DB   1 MPSELIPTLLLSGTFSSHATAMSIINSENEMDM----- --YEELWKHP 44
QY   95 KYVSMEENGFPIFYKNNNFIKTNSGFSVLEFMNFFGIISKFPMAPETGYIKDSK 154
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
QY   45 KVINILDEKEKPQVEKUNIGFDHNAUNNYTLGNKFAIILNKPYAWVLGPIDAK 104
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
QY   155 DEERVFYS 162
DB   1 I I I
QY   105 --RRVMKI 110
DB   1 I I I

RESULT          6
Q44DH7 PRELIMINARY; PRT: 365 AA.
ID Q94DH7 AC
DT 01-DEC-2001 (TREMBUREL 19, Created)
DI 01-DEC-2001 (TREMBUREL 19, Last sequence update)
DE 01-DEC-2001 (TREMBUREL 19, Last annotation update)
DEF PUTATIVE CYSTEINE PROTEINASE.
DCL GN P0518C01.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC Ehrharteaceae; Oryzoideae; Oryza.
NCBI_TaxID=4530;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=Cv. NIPPONBARE;
RA Susaki T.; Matsumoto T.; Yamamoto K.;
FE "Oryza sativa" diff. cultivars; J. Genom. DNA, Chromosome 1, PAC
EB submitted (Feb-2001) to the EMBL, Genbank, DDBJ databases.
DL EMBL: AP003277; BAB63672.1;
SR SEQUENCE 365 AA; 39927 MW; 4E39D55F9FE0BEFE CRC64;

Query Match              18.6%; Score 166.5; DB 10; Length 365;
Best Local Similarity    38.7%; Pred. No. 3.2e-05;
Matches 36, Conservative 35, Mismatches 31, Indels 7, Gaps 2;

QY   86 FEETPKKYHKVYSMEENGPFELYKNNMFIKTNSGFSVLEFMNFGDLSKEEFMAR 145
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
QY   52 PEKMAYPKAYSLEEKIFPEVFVDKNIINHIDEENKKITGYWLGLNEFADLTDFEKAA 111
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
QY   146 FTGVY----KSKDKDERVKFSRSVASSEPE 173
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
QY   112 YSLTILTPAFNSML QLFYYEVENASLPKE 142
DB   1 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I

RESULT          7
Q42673 PRELIMINARY; PRT: 367 AA.
ID Q42673 AC
DT 01-NOV-1996 (TREMBUREL 01, Created)
DI 01-NOV-1996 (TREMBUREL 01, Last sequence update)
DE 01-DEC-2001 (TREMBUREL 19, Last annotation update)
DEF PAPAYA PROTEINASE OMEGA (EC 3.4.22.6).
DCL GN OMEGAIL.
OS Carica papaya (Papaya)

```


1. The first part of the document discusses the importance of maintaining accurate records of all transactions and the role of the accounting system in providing reliable financial information. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various methods used to collect and analyze financial data, including the use of statistical techniques and the application of mathematical models. It highlights the importance of using appropriate methods to ensure the accuracy and reliability of the results.

3. The third part of the document discusses the challenges faced by organizations in managing their financial resources effectively. It identifies key areas such as budgeting, forecasting, and risk management, and provides strategies to address these challenges.

4. The fourth part of the document focuses on the role of the accounting system in providing timely and accurate financial information to management and stakeholders. It discusses the importance of maintaining up-to-date records and the use of technology to improve the efficiency of the accounting process.

5. The fifth part of the document discusses the importance of maintaining accurate records of all transactions and the role of the accounting system in providing reliable financial information. It emphasizes the need for transparency and accountability in financial reporting.

6. The sixth part of the document outlines the various methods used to collect and analyze financial data, including the use of statistical techniques and the application of mathematical models. It highlights the importance of using appropriate methods to ensure the accuracy and reliability of the results.

7. The seventh part of the document discusses the challenges faced by organizations in managing their financial resources effectively. It identifies key areas such as budgeting, forecasting, and risk management, and provides strategies to address these challenges.

8. The eighth part of the document focuses on the role of the accounting system in providing timely and accurate financial information to management and stakeholders. It discusses the importance of maintaining up-to-date records and the use of technology to improve the efficiency of the accounting process.

9. The ninth part of the document discusses the importance of maintaining accurate records of all transactions and the role of the accounting system in providing reliable financial information. It emphasizes the need for transparency and accountability in financial reporting.

10. The tenth part of the document outlines the various methods used to collect and analyze financial data, including the use of statistical techniques and the application of mathematical models. It highlights the importance of using appropriate methods to ensure the accuracy and reliability of the results.

Db 1 MAMIPISIKTLEVAICIFVYMGLS--FGDPSVIGSQNDLTSTPFIQLFESWMLKHKKI 58
 QY 97 YSSMEENQRFYKONNMFKITNSQGFYSYVLEMMNFPGDLSKEEFMARFTGYI 150
 Db 59 YKNIPEYTFEIKFKNLKYIDETNKKNSYWGILNVIADMSNDEFEKRYTGS 112

RESULT 5

US-08-847-171B 7
 : Sequence 7, Application US/08827171B
 : Patent No. 6254869
 : GENERAL INFORMATION:
 : APPLICANT: CAROLYN PETERSEN
 : APPLICANT: JIN-XING HUANG
 : TITLE OF INVENTION: CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
 : TITLE OF INVENTION: PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
 : TITLE OF INVENTION: TREATMENT, DIAGNOSTICS AND
 : TITLE OF INVENTION: DETECTION OF
 : TITLE OF INVENTION: CRYPTOSPORIDIUM PARVUM
 : NUMBER OF SEQUENCES: 16
 : CORRESPONDENCE ADDRESS:
 : ADDRESSES: PETERS, VERNY, JONFS & RIF A
 : STREET: 385 Sherman Avenue, Suite 6
 : CITY: Palo Alto
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94306-1840
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
 : COMPUTER: PC
 : OPERATING SYSTEM: WINDOWS
 : SOFTWARE: Wordperfect 6.0a WINDOWS
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US-08-847-171B
 : FILING DATE:
 : CLASSIFICATION: 536
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 62/011,233
 : FILING DATE: March 27, 1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Rand Verry
 : REGISTRATION NUMBER: 46,518
 : REFERENCE/DOCKET NUMBER: (HV)
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 324-1677
 : TELEFAX: (415) 324-1678
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 345 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : ORIGINAL SOURCE:
 : ORGANISM: Carica
 : US-08-847-171B 7

Query Match 17.4%; Score 157.5; DB 4; Length 345;
 Best Local Similarity 42.5%; Pred. No. 2.8e-08;
 Matches 27; Conservative 26; Mismatches 46; Indels 5; Gaps 2;

Db 1 MAMIPISIKTLEVAICIFVYMGLS--FGDPSVIGSQNDLTSTPFIQLFESWMLKHKKI 58
 QY 97 YSSMEENQRFYKONNMFKITNSQGFYSYVLEMMNFPGDLSKEEFMARFTGYI 150
 Db 59 YKNIPEYTFEIKFKNLKYIDETNKKNSYWGILNVIADMSNDEFEKRYTGS 112

RESULT 6

US-09-120-365-76
 : Sequence 76, Application US/09120365
 : Patent No. 6103514
 : GENERAL INFORMATION:
 : APPLICANT: Natori, Shunji
 : TITLE OF INVENTION: NEW PROTEASE
 : FILE REFERENCE: 32290-144749
 : CURRENT APPLICATION NUMBER: US/09/120,365
 : CURRENT FILING DATE: 1998-07-22
 : EARLIER APPLICATION NUMBER: JP 9-333 474
 : EARLIER FILING DATE: 1997-11-18
 : NUMBER OF SEQ ID NOS: 101
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 76
 : LENGTH: 380
 : TYPE: PRT
 : ORGANISM: Actinidin
 : US-09-120-365-76

Query Match 15.3%; Score 137; DB 3; Length 380;
 Best Local Similarity 26.9%; Pred. No. 3.9e-06;
 Matches 32; Conservative 31; Mismatches 42; Indels 14; Gaps 3;

QY 36 KNIIIAITIAIFIVLVTVSYIITNTSDKIDDFVFGDYVDPATREYRKSFEFKKKYHK 95
 Db 5 KSFVMSLFFSTLLILSLA-FNAKNLTORTND-----EVKAMYESWLKYGK 51
 QY 96 VYSSMEENQRFYKONNMFKITNSQ-GFSYVLEMMNFHDLKKEEFMARFTGYIKDS 153
 Db 52 SYNLSGEWEPPEIKETLFFIDFHNATNPYSKYKVLNQFADITDERFPSTYIAFTSGS 110

RESULT 7

US-09-515-039-76
 : Sequence 76, Application US/09515039
 : Patent No. 6214599
 : GENERAL INFORMATION:
 : APPLICANT: Natori, Shunji
 : TITLE OF INVENTION: NEW PROTEASE
 : FILE REFERENCE: 32290-144749
 : CURRENT APPLICATION NUMBER: US/09/515,039
 : CURRENT FILING DATE: 2000-03-06
 : EARLIER APPLICATION NUMBER: JP 9-333 474
 : EARLIER FILING DATE: 1997-11-18
 : NUMBER OF SEQ ID NOS: 101
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 76
 : LENGTH: 380
 : TYPE: PRT
 : ORGANISM: Actinidin
 : US-09-515-039-76

Query Match 15.3%; Score 137; DB 4; Length 380;
 Best Local Similarity 26.9%; Pred. No. 3.9e-06;
 Matches 32; Conservative 31; Mismatches 42; Indels 14; Gaps 3;

QY 36 KNIIIAITIAIFIVLVTVSYIITNTSDKIDDFVFGDYVDPATREYRKSFEFKKKYHK 95
 Db 5 KSFVMSLFFSTLLILSLA-FNAKNLTORTND-----EVKAMYESWLKYGK 51
 QY 96 VYSSMEENQRFYKONNMFKITNSQ-GFSYVLEMMNFHDLKKEEFMARFTGYIKDS 153
 Db 52 SYNLSGEWEPPEIKETLFFIDFHNATNPYSKYKVLNQFADITDERFPSTYIAFTSGS 110

RESULT 8

US-08-860-255A-5
 : Sequence 5, Application US/08860255A
 : Patent No. 6274336
 : GENERAL INFORMATION:
 : APPLICANT: Abdel-Mequid, Sherin


```

: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 05/08/404,445
: FILING DATE: 15-MAR-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 68961/1994
: FILING DATE: 15-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 23296/1995
: FILING DATE: 19-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 4508-002-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 323 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-404-445-3

Query Match 15.0%; Score 134.5; DB 1; Length 323;
Best Local Similarity 34.0%; Pred. No. 5.6e-06;
Matches 36; Conservative 19; Mismatches 30; Indels 21; Gaps 3;

QY 43 TIAPIVAVVTVSLYITNTSDVIDDFGVGVVGRATPVPKSFPEFYKYVYKVSMMFF 162
: 1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
: 5 IYLVFVAVVKSAY-----DPLKAPNY-----PEEFVPEKKNYSSEVE 44

QY 103 PQQFPEYKQNMFKTTNSQGSYVLPNPFEDISKEEFMAFFTG 148
: 1111111111111111111111111111111111111111111111111111111
: 45 KLEPFKIQNFIINFNQNSAKYVFNKFSKSLSKDTIAKYG 89

DB 45 KLEPFKIQNFIINFNQNSAKYVFNKFSKSLSKDTIAKYG 89

RESULT 11
US-09-500-651-2
: Sequence 2, Application US/09500651
: GENERAL INFORMATION:
: APPLICANT: ASANO, MINAO
: APPLICANT: KAWAI, MISAKO
: APPLICANT: MIWA, TETSUYA
: APPLICANT: NIO, NORIKI
: TITLE OF INVENTION: AMINOPEPTIDASE GX, AND A METHOD OF
: TITLE OF INVENTION: HYDROLYSING A PROTEIN WITH THE SAME
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 05/08/404,445

```

```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/813,591
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 030458/1997
: FILING DATE: 14-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 10-845-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 351 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-500-651-2

Query Match 15.0%; Score 134.5; DB 1; Length 351;
Best Local Similarity 37.7%; Pred. No. 6.3e-06;
Matches 26; Conservative 19; Mismatches 23; Indels 1; Gaps 1;

QY 81 EYKSFEEFKKKYKVVSSMERENQFFYKQNMFKTTNS-QGFSYVLPNPFEDISK 139
: 1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
: 20 ELMSYKQWLKVKYVNALGKFKPKFKKQIFKQIFKQIFKQIFKQIFKQIFKQIFKQIF 79

DB 20 ELMSYKQWLKVKYVNALGKFKPKFKKQIFKQIFKQIFKQIFKQIFKQIFKQIFKQIF 79

QY 140 EEFMAFFTG 148
: 1111111111111111111111111111111111111111111111111111111
: 80 EEFMAFFTG 88

RESULT 12
US-09-500-651-2
: Sequence 2, Application US/08813591
: Patent No. 5824534
: GENERAL INFORMATION:
: APPLICANT: ASANO, MINAO
: APPLICANT: KAWAI, MISAKO
: APPLICANT: MIWA, TETSUYA
: APPLICANT: NIO, NORIKI
: TITLE OF INVENTION: AMINOPEPTIDASE GX, AND A METHOD OF
: TITLE OF INVENTION: HYDROLYSING A PROTEIN WITH THE SAME
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 05/08/813,591
: FILING DATE: 07-MAR-1997
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 051848/1996
: FILING DATE: 08-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 030458/1997
: FILING DATE: 14-FEB-1997
: ATTORNEY/AGENT INFORMATION:

```


OY 144 ARPTG 148
| |
Db 119 ATRLG 123

Search completed: October 9, 2002, 08:25:01
Job time: 153 sec

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in financial matters. The text outlines various methods for organizing and storing data, including digital databases and physical filing systems. It also mentions the need for regular audits and reviews to ensure the integrity of the information.

2. The second section focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides guidelines for effective communication, such as using appropriate language, listening actively, and providing feedback. It also discusses the benefits of open communication and how it can foster a collaborative work environment.

3. The third part of the document addresses the challenges of managing resources and time. It acknowledges that resources are often limited and that time is a precious commodity. The text offers strategies for prioritizing tasks, delegating responsibilities, and managing time effectively. It also discusses the importance of monitoring resource usage and making adjustments as needed to ensure that the organization is operating efficiently.

4. The final section discusses the importance of continuous improvement and innovation. It emphasizes that organizations must be willing to embrace change and seek out new ways to improve their processes and products. The text provides examples of successful innovation and discusses the factors that contribute to a culture of continuous improvement. It also mentions the importance of staying up-to-date with industry trends and technologies.

GenCore version 4.5
Copyright (c) 1993 - 2000 GenCore Inc.

CM protein - protein search, using sw model

Run on: October 9, 2002, 08:25:03, Search time: 06:52 seconds
(without alignments)
160.697 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 2141

Sequence: 1 MDJGNVHEDYISCPYA KKCTGCTIVHPVVPVINSI 401

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query %	Match Length	DB ID	Description
1	605	28.3	462	1 P021_APATH	P43297 arabidopsis
2	586.5	27.4	348	1 P4P3_CARPA	P10056 carica papa
3	586.5	27.1	362	1 CYSF_HAVS	P25803 phascolus v
4	581.5	27.2	362	1 CYSF_VTGM	P12412 vigna mungo
5	581	27.1	360	1 CYSF_HMSF	P43156 hemerocalli
6	577.5	27.0	344	1 CATL_MOUSE	P06797 mus musculu
7	576.5	26.9	380	1 ACTN_ACTCH	P00785 actinidia c
8	576.5	26.9	458	1 OPVA_OVISA	P25776 oryza sativ
9	574	26.8	345	1 ANAN_ANACO	P00884 ananas com
10	568.5	26.6	334	1 CATL_RAT	P07154 rattus norv
11	567.5	26.5	352	1 P4P2_CARPA	P14080 carica papa
12	562	26.2	476	1 CYS2_DITH1	P44984 dictyosteli
13	560	26.2	334	1 C3L2_HUMAN	P06711 homo sapien
14	558	26.1	321	1 CYS3_HOMAM	P25784 homarus ame
15	557.5	26.0	323	1 CATL_HUMAN	P07311 homo sapien
16	552	25.8	334	1 CATL_PIG	P28244 sus scrofa
17	547.5	25.7	471	1 OYR_OVISA	P25777 oryza sativ
18	545.5	25.6	428	1 CYS4_BEANA	P25251 brassica na
19	541.5	25.3	471	1 CYS1_HORVU	P25249 hordeum vul
20	541.5	25.3	373	1 CYS2_HORVU	P25250 hordeum vul
21	541	25.3	329	1 CATK_MACFA	P27641 macaca fasc
22	540	25.2	331	1 CYS5_HUMAN	P25774 homo sapien
23	539	25.2	323	1 CATK_RABIT	P43236 cryctolagus
24	539	25.2	340	1 CATS_MOUSE	P07070 mus musculu
25	538	25.1	330	1 CATK_PIG	P04133 sus scrofa
26	537	25.1	334	1 CATL_BOVIN	P25975 bos taurus
27	535	25.0	329	1 CATK_HUMAN	P43235 homo sapien
28	528	24.7	329	1 CATK_RAT	P05186 rattus norv
29	526	24.6	345	1 P4P4_CARPA	P06784 carica papa
30	514.5	24.3	448	1 P4P4_CAPPA	P05994 carica papa
31	518.5	24.2	422	1 CYS1_HOMAM	P13277 homarus ame
32	515	24.1	423	1 CATK_MOUSE	P25697 mus musculu
33	513	24.0	433	1 TEST_RAT	P15242 rattus norv

34	505	23.6	323	1 CYS2_HOMAM	P25782 homarus ame
35	503.5	23.5	344	1 CYS5_DICDI	P46640 dictyosteli
36	498	23.3	442	1 CYS4_DITH1	P44983 dictyosteli
37	497.5	23.2	356	1 CYS3_LYCES	P40143 lycopersico
38	496.5	23.2	360	1 CYS2_MAIZE	P10717 zea mays (m
39	495	23.1	343	1 CYS1_DICDI	P04984 dictyosteli
40	493.5	23.0	569	1 CYSF_PLAFA	P25805 plasmodium
41	492	22.9	506	1 CYSF_PAVN	P46102 plasmodium
42	490.5	22.9	330	1 CATS_PAT	P22765 rattus norv
43	487	22.7	362	1 ALPH_HORVU	P05167 hordeum vul
44	482.5	22.5	341	1 CATL_DROME	P25029 drosophila
45	481	22.5	333	1 CATJ_MOUSE	P09014 mus musculu

ALIGNMENTS

RESULT	1
RD21_ARATH	
ID	RD21_APATH STANDARD PFT: 462 AA
AC	P43297
DT	01-NOV-1995 (rel. 32, Created)
DT	01-NOV-1995 (rel. 32, Last sequence update)
DT	01 MAR 2002 (rel. 41, Last annotation update)
DE	Cysteine proteinase PPTA precursor (pr 3 4 22..)
GN	RD21A OR A1047128 OR P2C19.31
OS	Arabidopsis thaliana (Mouse-ear cross).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons; Rosidae;
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV, COLUMBIA;
RX	MEDLINE=93314960; PubMed=8325504;
EA	Koizumi M., Yamaguchi-Shinozaki K., Tsuji H., Shinozaki K.;
RI	"Structure and expression of two genes that encode distinct drought-
RI	inducible cysteine proteinases in Arabidopsis thaliana.";
RL	Gene 129:175-182(1993).
RN	[2]
RF	SEQUENCE FROM N.A.
RC	STRAIN=CV, COLUMBIA;
RX	MEDLINE=2103719, PubMed 1130712,
EA	Theologis A., Ecker J.P., Palm C.J., Federspiel N.A., Kaul S.,
EA	White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
EA	Buchler E., Chan A., Chen Q., Chen H., Cheuk P.F., Chin C.W.,
EA	Cheng M.K., Chen L., Conway A.R., Conway A.P., Creasy T.H., Dewar K.,
EA	Duan P., Egan P., Feldblyum T.V., Feng J.-D., Feng B., Fujii C.Y.,
EA	Gill T.E., Godwin A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
EA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
EA	Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
EA	Lander-Hogrefe S., Lee A., Lee J.M., Lee C.A., Li J.H., Li Y.-P.,
EA	Lin X., Liu X., Liu Z.A., Luo J.S., Maill K., Matzli A.,
EA	Muller J., Miranda M., Nguyen M., Nierman W.C., Osborne B.J.,
EA	Pai C., Peterson J., Pham P.K., Rhee M., Rooney T., Sealey D.,
EA	Sakane H., Salas S.L., Schwartz J.F., Shinn P., Southwick A.M.,
EA	Sao H., Talbot L., Tamminga G., Tatum M., Tawn E.D.,
EA	Teeterback T., Van Aken G., Vaynsberg M., Vysotskaya V.S., Walker M.,
EA	Wu S., Xu G., Fraser G.M., Venter J.C., Davis R.W.,
EA	"Sequence and analysis of chromosome 1 of the plant Arabidopsis
EA	thaliana.";
CC	Nature 408:818-823(2000).
CC	INTRODUCTION. BY HIGH SALT CONDITIONS.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC	PAPAIN FAMILY OF THIOL PROTEASES.
CC	-----
CC	this SWISS PRO entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).

DR	PROSITE; PS00639; THIOL-PROTEASE-HIS; 1.
DR	PROSITE; PS00640; THIOL-PROTEASE-ASN; 1.
KW	HYDROLASE; Thiol protease; Zymogen; Signal; 3D-structure.
FT	SIGNAL 1 16 POTENTIAL.
FT	PROPEP 17 132 ACTIVATION PEPTIDE.
FT	CHAIN 133 348 CARICAIN.

Query Match	27.4%	Score=586.5	PR 1,	Length 348;
Best Local Similarity	36.6%	Prod No 18e-35;		
Matches 135,	Conservative 64,	Mismatches 139,	Indels 31,	Gaps 12,

[illegible]

DR PROSITE: PS00139; THIOLESTERASE_CYS; 1.
DR PROSITE: PS00649; THIOLESTERASE_HIS; 1.
DR PROSITE: PS00640; THIOLESTERASE_ASN; 1.
KW Hydrophobic; Thiol protease; Signal; Allergen; 3D-structure.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 1 126 ACTIVATION PEPTIDE.
FT CHAIN 127 380 ACTINIDAIN.
FT ACT_SITE 151 151
FT ACT_SITE 288 288
FT ACT_SITE 308 308
FT DISULFID 148 191
FT DISULFID 182 224
FT DISULFID 282 332
FT CONFLICT 96 96 D -> G (IN REF. 2).
FT CONFLICT 108 108 E2 -> V (IN REF. 2).
FT CONFLICT 123 124 E2 -> V (IN REF. 2).
FT CONFLICT 167 170 VTGV -> TSGS (IN REF. 3).
FT CONFLICT 181 181 E -> G (IN REF. 3).
FT CONFLICT 184 184 R -> G (IN REF. 2).
FT CONFLICT 192 192 N -> D (IN REF. 3).
FT CONFLICT 206 206 N -> D (IN REF. 3).
FT CONFLICT 212 212 P -> G (IN REF. 3).
FT CONFLICT 223 227 E2MD -> DQVA (IN REF. 3).
FT CONFLICT 230 231 NE -> EQ (IN REF. 3).
FT CONFLICT 240 240 E -> G (IN REF. 2).
FT CONFLICT 272 274 HVS -> GVA (IN REF. 3).
FT CONFLICT 286 286 I -> V (IN REF. 3).
FT CONFLICT 290 291 VT -> V (IN REF. 3).
FT CONFLICT 301 301 I -> V (IN REF. 3).
FT CONFLICT 307 307 K -> F (IN REF. 2).
FT CONFLICT 349 349 H -> Y (IN REF. 2).
FT CONFLICT 360 360 P -> S (IN REF. 2).
FT CONFLICT 371 380 GVDVGRYSA -> E (IN REF. 2).
FT STRAND 131 132
FT HELIX 133 146
FT TURN 137 137
FT STRAND 144 144
FT TURN 146 147
FT HELIX 151 168
FT STRAND 174 174
FT HELIX 176 182
FT STRAND 185 185
FT TURN 186 187
FT STRAND 188 188
FT HELIX 190 192
FT HELIX 196 206
FT TURN 207 207
FT STRAND 209 209
FT STRAND 211 211
FT TURN 212 214
FT HELIX 226 230
FT STRAND 234 234
FT STRAND 238 241
FT TURN 243 245
FT HELIX 247 256
FT STRAND 259 263
FT HELIX 268 272
FT STRAND 277 278
FT STRAND 288 298
FT TURN 299 300
FT STRAND 301 307
FT STRAND 310 310
FT TURN 312 313
FT TURN 315 315
FT STRAND 316 316
FT TURN 317 318
FT STRAND 319 323
FT TURN 328 329
FT HELIX 331 333
FT TURN 334 335
FT STRAND 339 342
SQ SEQUENCE 380 AA; 421 ER MW; 70FR1A224537324 CRG64.

Query Match 26.9%; Score 576.5; DB 1; Length 380;
Best local Similarity 35.3%; Pred. No. 1e-34;
Matches 129, Conservative 73, Mismatches 132, Indels 31, Gaps 11;
QY 36 KNIITATITATITIVIVTVTS; YITNTTSKTIQFVPCWVIDPATPFYVSPFPFKKKYHK 95
DB 5 KSPVSNLLFFSTLLIISLA-ENAKNLQRTND-----EKKAMYESWLKYGK 51
QY 96 VYSSMEENQRPPIYKONNNFKITNSQ-GFSYVLEMMFPGDLSKEEFMAFPTGYIKDSK 154
DB 52 SYNSDGEWKRREIEKTEFLFIDEHNAUTINESYKVLGNQFADLTDEEFISYLDGFISS- 110
QY 155 LGEFVFKSSVSAASESEEP--VINEINWVAGVNI-LNKKNVSSWAFSAVALEGA 212
DB 111 -- NKIKVS-NPEYEPFGVLFYSYVWESAGAVLIKSGESGGSWASATATVEST 163
QY 213 PPAQNSP-RISSESGFVLSKQNF-SQSTPMGLAFQYATPKYKLTINDYFYFAEEK 272
DB 164 NKIVTC-VLISLSPQPIIDQVGRFQNKNGYITDQGFLLNNGJINTEENYETAGDG 222
QY 273 TC MGPSTENYTFIPVKAYKYVFFENINAKLTALAKYGEISVAZGAPCTFFQYKSGVFD 331
DB 223 EGNLDQNEKYVDT DYERVVYNNWALQAVT-YGVSVVADAGAEAKHYSSGIFT 279
QY 332 APGKIKVNGVWVLEIMDEINKYVWLVNSGEAWSEKGYKIKIALHSGKKGTGGIIVE 391
DB 280 GPSTGTAIDVAVTIVGCTEGT -DYVKNKSWPTTWFFCYMP(IIPNVGGAGTCGATM 337
QY 392 PVYPV 396
DB 338 PSYPV 342
RESULT 8
ID ORYA_ORYSA STANDARD; PRT; 458 AA.
AC P25776;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Oryza alpha chain precursor (EC 3.4.22.-).
OS Oryza sativa (rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Pharbitideae; Oryzoaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 129-135.
RC STRAIN-CV, NIPPONBARE; TISSUE=Seed;
RX MEDLINE=91358494; PubMed=1885617;
KA Watanabe H., Abe K., Emori Y., Hosoyama H., Arai S.;
ET "Molecular cloning and gibberellin-induced expression of multiple
ET cysteine proteinases of rice seeds (oryzaeins).";
RL J. Biol. Chem. 266:16897-16902(1991).
CC -P- TISSUE SPECIFICITY: EXPRESSED ONLY IN SEEDS.
CC -P- INDUCTION: BY GIBBERELLIC ACID (GA).
CC ! SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1, ALSO KNOWN AS THE
CC PAPAIN FAMILY OF THIOLESTERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D90406; PAA14402.1;
DR PIR: J00388; KHRZOA.
DR BSSP: P00785; 2ACT.
DR MRPSP: C01.029.
DR InterPro: IPR000118; Granulin.

Db 261 FQHYKPGVTPGCTPI NHAIVITGYQDQ-SSEKKEWVPSWCAWGGGRTGYIPIAPVVS 319

QY 381 GKKTGCGILVERPTVI 397

Db 420 SSFGLGCIAMEPLYPTL 446

RESULT 10

CATL_RAT

AC P07134; Q9QW07; STANDARD; PRT; 334 AA.

DT 01-APR-1988 (Rel. 07, Created)

DI 01-FEB-1991 (Rel. 17, Last sequence update)

DI 16-OCT-2001 (Rel. 40, Last annotation update)

DE Cathepsin L precursor (pI 3.4-22.15) (Major secreted protein) (Mrp)

DE (Cyclic protein-2) (CP-2).

GN CRL.

OS Rattus norvegicus (Rat).

OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,

OC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Pteropus

OX NCBI_TaxID=101116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=WISTAR; TISSUE=Kidney;

RX MEDLINE=88040047; PubMed=3666143;

RA Ishidoh K., Towatari T., Imajoh S., Kawasaki H., Kominami E.,

KA Katsumata N., Suzuki K.,

RT "Molecular cloning and sequencing of cDNA for rat cathepsin L.",

RL FEBS Lett. 223:69-73(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=9002543; PubMed=260413;

PA Ishidoh K., Kominami E., Suzuki K., Katsumata N.,

RT "Gene structure and 5' upstream sequence of rat cathepsin L.",

RL FEBS Lett. 259:71-74(1989).

RN [3]

RP SEQUENCE OF 1-42 FROM N.A. TISSUE SPECIFICITY, AND INDUCTION

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=21255611; PubMed=11456678;

RA Zabludoff S.D., Clifton M., Zabludoff J.N., Shoukova N., Wright W.W.,

RT "Male germ cells regulate transcription of the cathepsin L gene by rat

RT Sertoli cells".

RL Pediatr. 142:2338-2347(2001)

RN [4]

RP SEQUENCE OF 88-134 FROM N.A.

RC TISSUE=Sertoli cells;

RX MEDLINE=92168015; PubMed=1791930;

RA Erickson-Lawrence M., Zabludoff S.D., Wright W.W.,

RT "Cyclic protein 2, a secretory product of rat Sertoli cells, is the

RT proenzyme form of cathepsin L.",

RL Mol. Endocrinol. 5:1789-1798(1991).

RN [5]

RP SEQUENCE OF 18-37, FUNCTION, AND SUBCELLULAR LOCATION

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Sertoli cells;

RX MEDLINE=95236691; PubMed=7717858;

PA Boujrad N., Oywadeba S.B., Carrier M., Leclerc H., Martin R.M.,

PA Papadopoulos V.,

RA "Identification of a stimulator of steroid hormone synthesis isolated

RT from testis.",

RL Science 268:1629-1632(1995).

RN [6]

RP SEQUENCE OF 18-28, AND TISSUE SPECIFICITY.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Epidermis;

RX MEDLINE=2014186; PubMed=10699763;

PA Kawada A., Hara K., Kominami F., Tezuka T., Takahashi M., Takahara H.,

RT "Precursor of rat epidermal cathepsin L: purification and

RT immunohistochemical localization",

RL J. Dermatol. Sci. 23:36-45(2000)

RN [7]

RP SEQUENCE OF 114-288 AND 291-334.

RC TISSUE=Liver;

RX MEDLINE=88296890; PubMed=3402618;

RA Towatari T., Katsumata N.,

RT "Amino acid sequence of rat liver cathepsin L.",

RL FEBS Lett. 236:57-61(1988).

CC FUNCTION: Important for the overall degradation of proteins in

CC lysosomes. Procathepsin L is required for maximal stimulation

CC of steroidogenesis by TIMP1

CC CATALYTIC ACTIVITY: Specificity close to that of papain. As

CC compared to cathepsin B, cathepsin L exhibits higher activity

CC towards protein substrates, but has little activity on Z-Arg-Arg-

CC NHMeC, and no peptidyl-dipeptidase activity.

CC SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE

CC BONDS.

CC SUBCELLULAR LOCATION: Procathepsin L is secreted. Mature cathepsin

CC L is lysosomal.

CC TISSUE SPECIFICITY: Both mature cathepsin L and procathepsin L are

CC found in the upper epidermis. The lower epidermis predominantly

CC contains procathepsin L. In seminiferous tubules expression is

CC greater at stages VI-VII than at stages IX-XII.

CC INDUCTION: Expression in Sertoli cells is repressed by germ cells.

CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE

CC PAPAIN FAMILY OF THIOL PROTEASES

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>)

CC or send an email to license@sib.sib.ch

EMBL: Y00697; CA68691.1; ...

EMBL: AF025476; AAB81616.1; ...

EMBL: S85184; AAB21516.1; ...

PIF: S07098, KIPT.

HSP: P07711; ICJL.

MEROPS: C01.032; ...

InterPro: IPR000068; Peptidase_C1.

InterPro: IPR001659; ThiolPro_Lact_site.

PIF: PF00112; Peptidase_C1; 1.

PRINTS: PF02705, PAPA1.

PROSITE: PS00139; THIOL_PROTEASE_CYS; 1

PROSITE: PS00634; THIOL_PROTEASE_HIS; 1.

PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.

Hydrolase, Thiol Protease, Glycylproline, Lysosome, Tyrosine, Signal

FT SIGNAL 1 17

FT PROPEP 18 113 ACTIVATION PEPTIDE.

FT CHAIN 114 288 CATHEPSIN L, HEAVY CHAIN.

FT CHAIN 291 334 CATHEPSIN L, LIGHT CHAIN.

FT ACT_SITE 138 138 BY SIMILARITY.

FT ACT_SITE 276 276 BY SIMILARITY.

FT ACT_SITE 300 300 BY SIMILARITY.

FT DISULFID 145 178 BY SIMILARITY.

FT DISULFID 169 211 BY SIMILARITY.

FT DISULFID 259 322 INTERCHAIN (BY SIMILARITY)

FT CARBOHYD 221 221 N-LINKED (GLCNAC)

FT CONFLICT 238 238 A ... P (IN REF 1)

SYNOPSIS 334 AA: 4760 MW; AFAV475824APK (PC64;

Query March 26, 2001; Score 568.5; IP: 1; Length 334;

Best Local Similarity 37.28; Pred. No. 3.4e-34;

Matches 128; Conservative 62; Mismatches 105; Indels 49; Gaps 14;

QY 78 ATREYKRSF-----EPPKKYHKVYSMEEDENOFFEYIKNNMF1KTTNSQ-----GFSY 127

Db 17 ATKPFQOTNAQHQWQKTHPHLYGTNEEEWPP-AVWEKNRMICLHNGEYSNCKHGT- 74

QY 129 VLEWNPFGDLSKFFMAPFTGYIKDKDERVKSSRSVSASSEEEFVPNSINWVAGC 187

Db 75 -MEMNAPGDMTNEEFQIVNGYFHQKHKKGLRQDEPLM-----LOIPKIVDMREKGC 125

QY 188 VNPTRNKNKNGSCWAFSAVAALGATCAQTNKGLSLSHQVFDKSKNGNFGICGGTNG 247

Db 126 VTPVKNNGDGECSWAFSASGCLLEGQMF1KTK-LTISFQNVIVDSHPQNGNCGNGLMD 184

[illegible][illegible]

GenCore version: 4.5
Copyright (c) 1993 - 2000 - Computer Ltd

OM protein - protein search, using sw model

Run on: October 9, 2002, 09:22:29 : Search time 40.3 seconds
(without alignments)
949,060 Million cell updates/sec

Title: US 09 598-062 4

Perfect score: 2341

Sequence: 1 MDIGNNVEEHQETSGPYIA.....AKGIGLIVPEYVYVINGSL 491

Scoring table: ROSTM62

Gapop 10.0 , Gapext 0.5

Searched: 28138 seqs, 9608934 residues

Total number of hits satisfying chosen parameters: 28138

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	in	id	Description
1	633	29.6	364	2	T12039	cysteine proteinas
2	631	29.5	454	2	J03848	cysteine proteinas
3	630	29.4	364	2	T46430	cysteine proteinas
4	614.5	28.7	455	2	T12041	cysteine proteinas
5	614	28.7	374	2	T03941	cysteine proteinas
6	605	28.3	462	2	J03719	drought-inducible
7	603	28.2	454	2	S24632	cysteine proteinas
8	596.5	27.9	464	2	T06416	cysteine proteinas
9	594	27.7	352	2	T10501	fruit bromelain (E
10	592	27.7	361	2	T06708	cysteine proteinas
11	592	27.7	427	2	S57776	cysteine proteinas
12	593.5	27.6	361	2	S55634	caricain (EC 3.4.2
13	588.5	27.5	340	2	T10516	fruit bromelain (E
14	588.5	27.5	368	2	S47312	cysteine proteinas
15	588	27.5	324	2	T10518	fruit bromelain (E
16	587.5	27.4	355	2	T06122	cysteine proteinas
17	587.5	27.4	362	2	S22502	cysteine proteinas
18	586.5	27.4	348	2	T06433	caricain (EC 3.4.2
19	586	27.4	348	2	F84672	probable cysteine
20	585.5	27.3	356	2	T10514	probable stom brom
21	585	27.3	356	2	A86341	cysteine proteinas
22	584.5	27.2	325	1	S11581	cysteine proteinas
23	581	27.1	351	2	T10503	fruit bromelain (E
24	581	27.1	360	2	S57777	cysteine proteinas
25	580	27.1	326	2	S53027	cysteine proteinas
26	580	27.1	376	2	E85435	cysteine proteinas
27	577.5	27.0	334	1	K08352	cysteine proteinas
28	576	26.9	458	1	K08270	oryzain (EC 3.4.22
29	575.5	26.9	341	2	T45839	probable cysteine

ALIGNMENTS

RESULT 1

T12039

Cysteine proteinase (EC 3.4.22.1) ; precursor - kidney bean

C:Species: Phaseolus vulgaris (kidney bean)

C:Accession: T12039

R:Senyuk, V.; Becker, G.; Muentz, K.

submitted to the EMBL data library, October 1997

A:Description: Isolation of cDNA clone encoding cysteine proteinase (Cpl) from a coty

A:Reference number: Z17385

A:Accession: T12039

A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-364 <SEN>

A:Cross references: EMBL, Z99952

A:Experimental source: Cultivar Meljavan, cotyledon, clone ep6a

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

EC:3.4.22.1; domain: signal sequence; status: predicted <SIG>

EC:3.4.22.1; domain: propionin; status: predicted <PRO>

EC:3.4.22.1; domain: cysteine proteinase; status: predicted <CAT>

EC:3.4.22.1; Active Site: Cys, His, Asp; status: predicted

Query Match 29.6%, Score 633, DB 2, Length 364,

Best Local Similarity 36.8%, Pred. No. 1.9e+39,

Matches 138, Conservative 74, Mismatches 129, Indels 34, Gaps 11,

QY 35 LKNTIATLAIPIVLVTVISLYITNNTSKIDDFVPGDYDPATREYRKSEFEKKYH 94

DB 1 MCEMLTPTLLLSFTSHATAMS:INYSNVNMDM.....-VEEVLVKEP 44

QY 95 KYSSNPFENPHPEIVKQNNIKTNSQPSYVLEMPFQISKEPFMAFTGYIKDSK 154

DB 45 KVTNGLLEKREKQVKKLNGEIQHNAQNNIITGLNKFADITNEEYFAMLTGRTDAK 104

QY 155 LDEKVFKESSRVG...AFSEDEEFVFNSTINWVEAFVNPINQKNGSGSWAFSAVALEG 211

DB 105 --EPVWKPTNIGHPYAYNSHQI--PVHYWMLKTAVPITKQCGNGSWAFSTVAVER 160

QY 212 AITAGNPNGLPSQPAIVLTKQNPNTSGIMELAPVAFPKKYLCTNFTVYFAEE 271

DB 161 INNVICVFVSLSEQLVCLQREYDE GENGGLMAYKQFIQNSGIDTFEDPYQGTG 218

QY 272 KLEMSFENY:ELIWKVKYVFEN:NAISIALAKY:GFSVA:GAQ:IFPFVKSGVED 331

DB 219 GIC-DGTRKKIKVQIDGTFDVFNSNNENALKRAVS-HQFVSVATCASQALQLQSGVFT 276

QY 332 AFQGTQVNBGVVLYEIMDELINKEYLWLVFNSWGLAWGEKGYIKLA--LHEKKKGTGGL 389

DB 277 GKGTALDHCVVGVVGVCTENGV--DYWLVFNSWCTWCTGYFKMERNVPSTSEKCGIA 334

30	574	26.8	345	2	T07839	anandin (EC 3.4.22
31	570.5	26.6	356	2	T07840	anandin (EC 3.4.22
32	569.5	26.5	476	2	S4451	cysteine proteinas
33	568	26.5	467	2	T06529	cysteine proteinas
34	567.5	26.5	352	2	T09760	chymopapain (EC 3.
35	567.5	26.5	480	1	TAGH	actinidain (EC 3.4
36	566.5	26.5	334	1	KHRTL	cathpsin L (EC 3.
37	564.5	26.4	453	2	T07851	anandin (EC 3.4.22
38	564	26.4	443	2	T06198	cysteine proteinas
39	562	26.2	476	1	KRQEP	prestalk cathpsin
40	559.5	26.1	342	2	S71773	cysteine proteinas
41	559	26.1	313	2	S47433	cathpsin L (EC 3.
42	557.5	26.0	333	1	KHHL	cathpsin L (EC 3.
43	557	26.0	331	2	T06413	cysteine proteinas
44	557	26.0	480	2	T01207	cysteine proteinas
45	555	25.9	320	2	T19651	cysteine proteinas

A: Experimental source: cv. N67-10; nearly mature fruit
C: Superfamily: papain
C: Keywords: cysteine proteinase; hydrolase
F: 95-313/Product: fruit bromelain #status predicted <MAI>
F: 314-324/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F: 120,362,277/Active site: Cys, His, Asp #status predicted

```
Query Match      27.5%; Score 588; DB 2; Length 324;
Best Local Similarity 47.6%; Pred. No. 3,70-36;
Matches 120; Conservative 56; Mismatches 111; Indels 22; Gaps 10;

QY 84 KSPPEKFKYHYVSSMEENQRFETIKQMMPIETINSQCTRYVLFMHEFCHLSKHEF 142
DB 1111 311 311 311 311 311 311 311 311 311 311 311 311 311
DB 8 KRPEEMAEYGRIVKONDEKMRRFQTFKNNVKHIETFNRSRNGNSYTLGINTDMTKSEF 67
QY 143 MARPTGY---IKSKKLEDFVFKSRVSASESEEEFFVPFNSINWVLAGEVNPINOKNUGS 199
DB 1111 311 311 311 311 311 311 311 311 311 311 311 311
DB 68 VAQYTRVSLPINTFRPPVVSFDVNIISA-----VPGSLWPEYGAIVNEVENPQGS 119
QY 200 CWAFSAVALEGATCAOTNNGIPSLSEQPFVDCSKNGNFGCGGTMGIAFOYAIKNKYL 259
DB 1111 311 311 311 311 311 311 311 311 311 311 311 311
DB 120 CWAFAAIATVEGIYKIKTGY-IVSLSEQEVLLQAV---SYGKGGWVAKAYDFILSNNGV 175
QY 260 CTNDYPPFAEPTQMISECFNYTETPVAKYRYVFFFNINALKIALAKYGPISVAIQADQ 319
DB 1111 311 311 311 311 311 311 311 311 311 311 311 311
DB 176 TTEENYPYQAYQCTCANSEFPN--SAYITGYSYVRNDRSMYAVSNQ-PIAALIDASE 232
QY 320 TPQCFYKSGYFADAPCGTFKNHGCVLVVEYEMDEETIKRYWLVNPSWSEAWGEKGYIKLAIH 379
DB 1111 311 311 311 311 311 311 311 311 311 311 311 311
DB 233 N-FQYVNGVSGPGSTSLNHALITIGYQDSQSGTK-VWIVFNSWSSWGEGGYVPMAPG 290
QY 380 -SGKKGTGCTLVEPPVYPV1 397
DB 1111 311 311 311 311 311 311 311 311 311 311 311
DB 291 VSSSSGAGGIAMSPLEPTL 309
```

Search completed: October 9, 2002, 08:23:24
Job time: 56 sec

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in financial matters. The text outlines various methods for organizing and storing records, including digital databases and physical filing systems. It also mentions the need for regular audits and reviews to ensure the integrity of the data.

2. The second part of the document focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides guidelines for effective communication, such as using appropriate language, listening actively, and providing feedback. It also discusses the benefits of open communication, including improved collaboration and faster decision-making.

3. The third part of the document addresses the issue of resource management. It discusses the importance of identifying and allocating resources effectively to support the organization's mission. The text outlines various strategies for resource management, including budgeting, prioritization, and delegation. It also mentions the need for regular monitoring and evaluation of resource usage to ensure optimal performance.

4. The fourth part of the document discusses the importance of continuous learning and improvement. It emphasizes that organizations must stay up-to-date with the latest trends and technologies to remain competitive. The text outlines various methods for learning, such as training, conferences, and research. It also mentions the need for a culture of learning, where employees are encouraged to share knowledge and learn from each other.

5. The fifth part of the document discusses the importance of risk management. It outlines various strategies for identifying, assessing, and mitigating risks. The text emphasizes that risk management is an ongoing process that requires regular monitoring and updates. It also mentions the need for a risk management framework that is tailored to the organization's specific needs and goals.

6. The sixth part of the document discusses the importance of stakeholder engagement. It outlines various methods for identifying and engaging stakeholders, such as surveys, interviews, and focus groups. The text emphasizes that stakeholder engagement is essential for understanding the needs and expectations of different groups and for developing effective solutions. It also mentions the need for regular communication and reporting to stakeholders.

7. The seventh part of the document discusses the importance of innovation and creativity. It outlines various methods for fostering innovation, such as brainstorming, idea competitions, and cross-functional teams. The text emphasizes that innovation is essential for staying ahead of the competition and for developing new products and services. It also mentions the need for a culture of innovation, where employees are encouraged to think creatively and take risks.

8. The eighth part of the document discusses the importance of sustainability. It outlines various strategies for reducing the organization's environmental impact, such as energy conservation, waste reduction, and sustainable sourcing. The text emphasizes that sustainability is not just a moral imperative but also a business imperative, as it can lead to cost savings and improved reputation. It also mentions the need for regular monitoring and reporting of sustainability metrics.

9. The ninth part of the document discusses the importance of corporate social responsibility (CSR). It outlines various strategies for promoting social and environmental responsibility, such as community outreach, employee volunteerism, and ethical sourcing. The text emphasizes that CSR is essential for building a positive reputation and for contributing to the well-being of society. It also mentions the need for regular monitoring and reporting of CSR activities.

10. The tenth part of the document discusses the importance of legal and regulatory compliance. It outlines various strategies for ensuring that the organization is compliant with all applicable laws and regulations, such as regular audits, legal counsel, and employee training. The text emphasizes that compliance is essential for avoiding legal penalties and for maintaining the organization's integrity. It also mentions the need for a strong legal and regulatory framework that is tailored to the organization's specific needs and goals.

YY Petersen C, Huang J;
 XX WPI: 2001-080923/41;
 XX N-PSDB: AAF44094.
 XX New vaccines and cryptopain protein for diagnosis and treatment of
 XX Cryptosporidium species infection.
 XX Claim 4: Fig 4: 32pp: English.
 XX The invention relates to vaccines and cryptopain protein for diagnosis
 XX and treatment of cryptosporidium species infection. Cryptopain protein
 XX can be expressed by standard recombinant methodology. The vaccines
 XX comprising the cryptopain protein or its fragments are useful for active
 XX immunization of animals and humans against cryptosporidium infection, or
 XX for production of passive immune products in admixture with an adjuvant.
 XX Cryptopain can be used for prophylaxis, therapy, diagnosis, diagnosis and
 XX detection purposes. The present sequence represents the pre and pro
 XX regions of *C. parvum* cryptopain protein.
 XX Sequence 175 AA;
 XX
 Query Match 41.8%; Score 896; DB 22; Length 175;
 Best Local Similarity 100.0%; Pred. No. 6, 20-77;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 QY 1 MERNVVEHGVYVSSGYIALINGTQOREPKKLNIIATLIAIFVLVIVSVLYITN 60
 Db 1 MERNVVEHGVYVSSGYIALINGTQOREPKKLNIIATLIAIFVLVIVSVLYITN 60
 QY 61 NTSKIDMVPVGVVDVATREYKSEEPKVKYHKVYSSMEENQREYKNNMFIKTT 120
 Db 61 NTSKIDMVPVGVVDVATREYKSEEPKVKYHKVYSSMEENQREYKNNMFIKTT 120
 QY 121 NQGFSEVLENNFGHLSKEFWARTGYIKESKIEEVEPKSKVSASSEPEFV 175
 Db 121 NQGFSEVLENNFGHLSKEFWARTGYIKESKIEEVEPKSKVSASSEPEFV 175
 RESULT 4
 AAB65770
 ID AAB65770 standard; Protein: 396 AA.
 XX AAB65770;
 DT 27 MAR 2001 (first entry)
 XX Cysteine protease #11.
 XX Cell death modulator; programmed cell death; PCP; apoptosis;
 KW forestry plant.
 XX Pinus radiata.
 OS W0200075331-A1.
 XX P0
 XX 14 DEC 2000.
 XX 02 JUN 2000; 2000W0200075331.
 XX 04 JUN 1999; 9905-0125932.
 XX (GPNF-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX Flett H, Lasham A;
 XX WPI: 2001-061724/07.
 XX N-PSDB: AAF44797.
 XX Novel detector against cell death polynucleotide useful for modulating

PT programmed cell death pathway and specific development pathways in
 PT forestry plant -
 XX claim 22, Pages 100-110, 142pp, English
 XX The present invention relates to coding sequences (see AAF44740-F44840
 XX and AAF44843-F44844) and proteins (see AAF5734-B65814) involved in
 XX programmed cell death (PCD; apoptosis). The coding sequences and proteins
 XX of the present invention are useful for modulating a PCD or cell death
 XX pathway and various developmental pathways in a forestry plant, by
 XX stably incorporating one of the present coding sequences into the genome
 XX of the forestry plant, where the coding sequence provides a PCP pathway
 XX that is not present in a native form of the forestry plant.
 XX Sequence 396 AA;
 XX
 Query Match 28.8%; Score 617; DB 22; Length 396;
 Best Local Similarity 38.8%; Pred. No. 6, 1e-50;
 Matches 144; Conservative 61; Mismatches 135; Indels 31; Gaps 12;
 QY 39 IATLIAIFVLVIVSVLYITNNTSDKIDDFVPGVYVDPATFFYKSPFFPKKY---- 93
 Db 23 VMASVSKATLIIIFLATTIWTISANASDSPF----- TDEAKSEELIILYDKWA 73
 QY 94 --HKVYSSW--EEENQREYKNNMFIKTNSQGFSEVLENNFGHLSKEFWARTGY 149
 Db 74 IHRTRISLSDSEHAKRFEIKDNVYKIDSVNQKDPYKLGINKFTDISNEEFKAM---H 130
 QY 150 IKDSKIDDFVPGVYVDPATFFYKSPFFPKKY----EEEFVYVNSINWVAGVYVNPVKNVWAFSAV 206
 Db 131 MTRMEKIKSLTRDGRTHSGTMYNSDNLPESIDWRKGVAVNPVKNVWAFSAV 190
 QY 207 AAI PCATCAQTNPCTPSISEQFVPSKQNGNPGYVNGCTMSLAPYAIKNKYLTNDRYP 266
 Db 191 AVEGLISYKTK-LVLSSEGLVDCSKENA GNGGLMDSATGYIIDNGGKAEDEY 247
 QY 267 YFAEKTQMDSPCE-NYIEIPVKAYKYVPPRNINAKLTALAKYPTSVATQADPTPFOFY 325
 Db 248 YLQASEVSPSKYKPCATAATIDGFEYPAKNEKAEKAVG-HQPSVAIEASGKDFQY 306
 QY 326 KSGVFDAVCGKVNKGVLVEYDMEDTNKYEWLVNPSWGFAGWPKGYIKLAIH-SGKKG 384
 Db 307 SKGVITQEGTEIDHQVAVGVYKSPG-GINYVIVNSWGFAGWPKGYIKLAIH-SGKKG 365
 QY 385 TCGILVEPVYP 395
 Db 366 KQYIAMPSYYP 376
 RESULT 5
 AAW19541
 ID AAW19541 standard; Protein: 463 AA.
 XX AAW19541;
 XX 17-SEP-1997 (first entry)
 XX Soybean thiol protease D3-alpha.
 XX Soybean thiol protease, germinator, transformant, Escherichia coli,
 KW recombinant expression vector.
 XX Glycine max.
 OS
 FH Key Location/Qualifiers
 FH Peptide 1..131
 FT /label= Signal
 FT Protein 132..463
 FT /label= Thiol_protease_D3-alpha
 XX JP091218/0-A.
 XX

QY 17 PYIALINGNOOREPNKKLNILIAIFIVLVVTVSYITNNTSKIDDFVPGDYVD 76
Db 11 ppilflh-----stripsmaipataaaallifcavaavaaa-----rtdevmgmy-- 56
QY 77 PATREYKKSFEFKKKYKHVYSMEENQREPIYKONNMFIKTTNSQGFSYVLEMMERGD 136
Db 57 -----elwlaKhqkaynqlqererrteitrdnirivdehualursytlgmrtad 106
QY 137 LSKEEFMARPTGYIKDSKIDDERVPKSSPVASSESEERFVP-----PNSINWVEAGCVNP 190
Db 107 lthceyrawylhtisqfm--rrvakaarasg-----ryyprpddmlpaavdgrtrgavnk 159
QY 151 IRNCKNNGSWAFSAVAALFVATCAATNKGTPSISFAGFVWFSKONNMFCTWZFMGLAP 250
Db 160 wKdHiciswafstiaureliqitgtcfhslseqelddcdray dajcngg'ldmyul 217
QY 251 QYAIKNYVICTNDYVPYFEKTKOMSPFENYIEIPVAYKYVFFPENINAKTALAKYGP 310
Db 218 qfildbqfildbqsy'rtqatn daskonakvstlgyedlqf9kzatkakabqj 275
QY 311 ISVAIOADOTPPQFYKSGVEDAPCGTKVNHGVVLVEYDMDEDTNKEYWLVHNSWCEANGE 370
Db 276 vetaicagardiqyosqftaccataldgvlacqyazqh jadywlvtruswslwqe 333
QY 371 KGVIKLA--LHSGKKGTGILVPEVPV 396
Db 334 sgyikwernlannytgkculamcasypv 361

Search completed: October 9, 2002, 08:31:26
Job time: 538 sec



GenCore version 4.5

Copyright (c) 1993 - 2000 Computer Ltd.

OM protein - protein search using sw model

Run on: Oct-09 10:51:06 2002 Search time: 40s 748 seconds
(without alignments)
171.825 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 2141

Sequence: 1 MGLNNVSEHGVISEPYIA.....EKSLQIVVFWIVNSI 401

Scoring table: BL0SUM62

Gapop 10.0 Gapext 0.5

Searched: 56222 seqs 17994999 residues

Total number of hits satisfying chosen parameters: 56222

Minimum DB seq length: 0

Maximum DB seq length: 2090000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREML_19:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_ordanelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2141	100.0	401	5 Q9X1T4	Q9X1T4 cryptospori
2	768.5	35.9	394	6 Q9X526	Q9X526 sarcocystis
3	633	29.6	444	10 Q44431	Q44431 phaseolus v
4	631	29.5	454	10 Q40922	Q40922 pseudotsug4
5	630	29.4	364	10 Q41110	Q41110 phaseolus v
6	624.5	29.2	359	10 Q40987	Q40987 homococall
7	621	29.0	365	10 Q40987	Q40987 oryza sativ
8	614.5	28.7	455	10 Q24323	Q24323 phaseolus v
9	614	28.7	374	10 Q24137	Q24137 nicotiana t
10	613	28.6	340	5 Q9NH99	Q9NH99 stylonychia
11	606	28.3	462	10 Q94BX1	Q94BX1 arabidopsis
12	605	28.2	343	10 Q4PS02	Q4PS02 astragalus
13	604	28.2	434	10 Q4A005	Q4A005 ipomoea bat
14	603	28.2	464	10 Q41064	Q41064 pisum sativ
15	599.5	28.0	462	10 Q93XQ9	Q93XQ9 ipomoea bat
16	597.5	27.9	452	10 Q94PM8	Q94PM8 arabidopsis

17 596.5 27.9 361 10 Q9Q999
18 596.5 27.9 466 10 Q49877
19 594 27.7 352 10 Q24641
20 592 27.7 361 10 Q9S114
21 592 27.7 427 10 Q43423
22 590.5 27.5 357 10 Q45673
23 588.5 27.5 340 10 Q23800
24 588.5 27.5 468 10 Q41646
25 588 27.5 324 10 Q23801
26 587.5 27.4 355 10 Q65493
27 586 27.4 348 10 Q9Z0H7
28 586 27.4 358 10 Q94HK7
29 585.5 27.3 352 10 Q23799
30 585 27.3 462 10 Q9SVL2
31 584.5 27.3 461 10 Q9FSS0
32 584.5 27.3 465 10 Q9S161
33 582 27.2 376 10 Q94B09
34 581 27.1 351 10 Q23791
35 580 27.1 325 5 Q27760
36 580 27.1 376 10 Q24195
37 579.5 27.1 357 10 Q9XP80
38 578.5 27.0 328 5 Q27759
39 578 27.0 360 10 Q82708
40 577.5 27.0 334 11 Q91020
41 576.5 26.9 434 11 Q90600
42 576.5 26.9 445 10 Q48886
43 575.5 26.9 441 10 Q48615
44 575.5 26.9 462 10 Q41690
45 571.5 26.7 346 10 Q9FJ47

ALIGNMENTS

RESULT 1

Q9X1T4 PRELIMINARY: PRT: 401 AA.
ID Q9X1T4
AC Q9X1T4;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CRYPTOPAIN PRECURSOR.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa, Coccidia, Kimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NINC;
PA Petersen C.J., Huang J.-X.;
FT "Characterization and expression of the gene for cryptopain, a
PT Cryptosporidium parvum cysteine protease homologue";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF091366; AA042940.1;
DR HSP; P07711; ICL.
DR MEROPS; C01.0PA;
DP Interpro; IPR000668; Peptidase_C1.
DP Interpro; IPR000169; Thiolprot_act_site.
DR Pfam; PF00112; Peptidase_G1; 1.
DR PRINTS; PR07005; PAPA1N.
DR PROSITE; PS00642; THIOL_PROTEASE_ASN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
RW Hydrolase, S4003, 175 POTENTIAL.
FT SIGNAL 1
FT CHAIN 176 401 CRYPTOPAIN.
SQ SEQUENCE 401 AA. 45467 MW; 55440E75CA61B70 CRC64;

Query Match 100.0% Score 2141; DB 5; Length 401;
Best Local Similarity 100.0%; Pos. 7 4e-151;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

27 1 MGLNNVSEHGVISEPYIA.....EKSLQIVVFWIVNSI 401


```

DB 164 WASTVAASVINAIVG-NLPLSEKQPLDGP-NGN-NGSGLMDYAEVLAANGCH 321
QY 261 TNDIYPFAEKTGM-----DSFGNYIPIPVKAYKYVPPNINALKTALAKYGPISVA 314
DB 222 TERSYVIMLEETGFGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 280
QY 315 IDAQVTHPEFYKSVFPAAGCGKVNHGVLVEYDMEDTNKRYWLVNSWSGEANGKGYL 374
DB 281 TEASGRNPQYSGVTPGCGCTFELRGVTAVNG TASKGHYIIVKNSGSHWSEKGYI 339
QY 375 KLAHSCK-KGTGSHILVERYPWIN 398
DB 340 PMPRGCKKHGLOGINKMASYPKEN 364

RESULT 8
ID 024323 PRELIMINARY: PRT: 455 AA.
AC 024323;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSTEINE PROTEINASE PR000639
OS Phascolus vulgaris (Kidney bean) (French bean).
OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae,
OC eurosids I, Fabales, Fabaceae, Papilionaceae, Phaseolidae, Phaseolus.
OX NCBI_TaxID:3885;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CV, MOLDAVIAN; TISSUE=COTYLEDON;
RA Senyuk V., Becker C., Muenz K.;
RT "Isolation of cDNA clone encoding cysteine proteinase (CP) from a
RT cotyledon specific cDNA library of germinating kidney bean seeds".
RL Submitted (Jul-1997) to the EMBL/GenBank/DDBJ databases
DR EMBL: Z54954; GenBank: AF070761;
DR HSSP: P00785; ZACT.
DR ANU-ZDPAGE: 024323;
DR InterPro: IPR000118; Granulin.
DR InterPro: IPR000668; Peptidase_C1
DR InterPro: IPR001211; FLEA2.
DR InterPro: IPR00169; Thiolprot_act_site.
DR Pfam: PF00496; granulin; 1.
DR Pfam: PF00112; Peptidase_C1; 1
DR PRINTS: PR00705; PAPAIN.
DR SMART: SM00277; GRAN; 1.
DR PROSITE: PS00118; PA2_HIS; UNKNOWN_1.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Signal; Thiol protease.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 127 455 CYSTEINE PROTEINASE.
SQ SEQUENCE 455 AA: 50253 MW: 32955BA4B55FC8B CRC64;

```

```

Query Match 28.7%; Score 614.5; DB 10; Length 455;
Best Local Similarity 39.0%; Pred. No. 1,2e-37;
Matches 140; Conservative 65; Mismatches 111; Indels 27; Gaps 10;

QY 76 DPAT-----RYPKSFEEFKKRYKRYVSSMFEENQRPETVYKQNNFKITNSQPSVVI 141
DB 26 DKATWEDREVNLSYELWEVHGKFNALGKFKPEKFNKFNPFHGGNAENRYKRG 45
QY 132 NEFDLSKEEFMAFFTYIKISNIDKDFPVKSSVSVSASESEELVP-----ENSINWVEA 185
DB 86 NRPADLINEEYFAYTIG-----TRIGPNPSTPTPS-NPYAPPVGGELIPUSVLMWKE 136
QY 196 GYNVEIENKNGDSWAFSAVALEGATAGTNGCLPSSEAGPFVMSKNGNPGQDGGSI 245
DB 137 GAVVPVKDASGSGWAFSAIAGVFINKTVTG-PIIISLSEDFVTD-TSYNMGTNGSL 194

```

```

QY 246 MCLAFQYAIKKYKICTINDYDYPFAAEKTCMDSFCENYIEIPVKAYKYVPPNINALKTAL 305
DB 195 MDYAFETIKNGGIDSEEDYDKGVDTGC-DEYKNAKVWSIDGYEVNIVYDEIALKAV 253
QY 306 AKYGPISVAIQAQIPEQFYKSGVSDAPCGIKVNHGVVLVEYEMEDTNKRYWLVNSWSG 365
DB 254 ANQ-PVSVAVEGGGPFEPGLYSSGVFGPGCTALHGGVVAQGY--GIIHGHHFVLPVNSWG 310
QY 366 FAWFKGKYIKIA--LHSCKKGTGSHILVERYPWIN 398
DB 311 ADWQFEGVIRFPNIGNSPSKCKGIALTEPSYPI 343

RESULT 9
ID 024137 PRELIMINARY: PRT: 374 AA.
AC 024137;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSTEINE PROTEINASE PR000639
OS Nicotiana glauca (Common tobacco).
OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots,
OC Asteridae, euasterids I, Solanales, Solanaceae, Nicotiana.
OX NCBI_TaxID:4077;
RN [1]
SEQUENCE FROM N.A.
RA Becker C., Muenz K.;
RT "cDNA cloning of a CPRI-homologous proteinase from germinating tobacco
RT seeds."
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: Z94173; GenBank: AF163171;
DR HSSP: P00785; ZACT.
DR InterPro: IPR005658; Peptidase_C1.
DR InterPro: IPR000639; Thiolprot_act_site.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Signal; Thiol protease.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 139 374 CYSTEINE PROTEINASE.
SQ SEQUENCE 374 AA: 42307 MW: 15414AFF85BF07DA CRC64;

```

```

Query Match 28.7%; Score 614; DB 10; Length 374;
Best Local Similarity 37.7%; Pred. No. 1e-37;
Matches 142; Conservative 67; Mismatches 148; Indels 30; Gaps 13;

QY 36 KHIITATITAIPIVIVTVVSLVITNNTSKIDDFVPGVVDPAI-----PFPKSFEEFK 90
DB 3 KTIITLLEFSSLSYAIDM-----SIIDYKNNHYARKWLTQSDQVKKRYEMWL 54
QY 91 KRYHVVSSMEFFNPPTTYKNNMFKITNSQ-PSVVIEMNPFEDLSKFFEMAPPTGY 149
DB 55 AFHGAAYNALGKKEKPEIFKDNLPFIEGHNNSGNPTKYKGLNPGADLTNEEYDTMYLGT 114
QY 150 LKSKLEDFPVKSSAVS---ASESEEEFVPIINSINWVAGTVNIFPKNGNSGSAWESAV 206
DB 115 PSWAP--PFPVSKNPSQYFASPPNPF--LPHSVIWWPKPAPVATKNGSGSGSWAFSTV 170
QY 207 AALPEZAIQAQINPGIPLSPAGPFVMSKNGNPGQDGGIMSLAQYAIKNKYLTNDIYP 266
DB 171 AAVEFINIVTVE-MITLSEGLVITGVV-NSQVNGSLMAYFPFIIISNGMDTERHYP 228
QY 257 YFAEKTGMDSFCENYIEIPVKAYKYVPPNINALKTALAKYGPISVAIQADQTFPFQYK 326
DB 229 YPVSVEGSG-DEYKNAKVWSIDGYEVNIVYDEIALKAV-HQVVCVAIEASGAFQIYS 285
QY 327 SAVFLAPGSLKVNHCQVVLVEYDMEDTNKRYWLVNSWSGEANGKGYIKLAIHSCKK-G 384

```




231	3	0.7	4	7	AAP60670	Sequence of amino	604	3	0.7	4	14	AAR37129	Cleavable peptide
232	4	0.7	4	7	AAP61356	Sequence of peptid	385	3	0.7	4	14	AAR44238	Peptide found in t
233	4	0.7	4	7	AAP61359	Sequence of peptid	386	3	0.7	4	14	AAR44240	Inhibitor of signa
234	4	0.7	4	7	AAP61655	Sequence of peptid	407	3	0.7	4	15	AAR52607	147 region of CMA
235	4	0.7	4	7	AAP61660	Sequence of peptid	408	3	0.7	4	15	AAR55033	CD83 region of L41
236	4	0.7	4	7	AAP60729	Synthetic repeatin	409	3	0.7	4	15	AAR55019	CD83 region of L41
237	4	0.7	4	7	AAP60123	Sequence of pharma	410	3	0.7	4	15	AAR55021	CD83 region of L41
238	4	0.7	4	8	AAP71305	Peptide component	411	3	0.7	4	15	AAR54529	Platelet aggregati
239	4	0.7	4	8	AAP71309	Peptide component	412	3	0.7	4	15	AAR44716	Tetrapeptide prepa
240	4	0.7	4	8	AAP71311	Peptide component	413	3	0.7	4	15	AAR47186	Anticlotting Coar
241	4	0.7	4	8	AAP71434	Immunomodulator pe	414	3	0.7	4	15	AAR48289	ACE inhibitor pept
242	4	0.7	4	9	AAP80873	Block region #4 con	415	3	0.7	4	15	AAR48293	ACE inhibitor pept
243	4	0.7	4	9	AAP82543	Tetrapeptide inhib	416	3	0.7	4	15	AAR48294	ACE inhibitor pept
244	4	0.7	4	9	AAP82680	Thymopentin analog	417	3	0.7	4	15	AAR47790	PLII fusion protei
245	4	0.7	4	10	AAP91605	Motif useful in to	418	3	0.7	4	15	AAR48247	Endoplasmic reticu
246	4	0.7	4	10	AAP91609	Motif useful in to	419	3	0.7	4	15	AAP49765	Pharmacofluoresce
247	4	0.7	4	10	AAP91617	Motif useful in to	420	3	0.7	4	15	AAR44794	Sequence of peptid
248	4	0.7	4	10	AAP91623	Motif useful in to	421	3	0.7	4	15	AAR51488	Human p53 NLS site
249	4	0.7	4	10	AAP93312	Peptide amide with	422	3	0.7	4	15	AAR51489	Human p53 CK II si
250	4	0.7	4	10	AAP93595	Peptide amide with	423	3	0.7	4	15	AAR50125	Tyrosinase inhibit
251	4	0.7	4	10	AAP92290	Chlorophyll peptide	424	3	0.7	4	15	AAR45963	Human hippocampal
252	4	0.7	4	11	AAP93395	Human interleukin	425	3	0.7	4	15	AAR52962	Human p53 NLS req
253	4	0.7	4	12	AAR13257	Cytotoxic cell fac	426	3	0.7	4	15	AAR52963	Human p53 CK II si
254	4	0.7	4	12	AAR13140	Synthetic neuropep	427	3	0.7	4	15	AAR51120	Genetic hippocamp
255	4	0.7	4	12	AAR10142	Synthetic neuropep	428	3	0.7	4	15	AAR55336	Inhibitor of tyros
256	4	0.7	4	12	AAR19564	Pibouletotide red	429	3	0.7	4	15	AAR59892	2-amino-6,7-dihydr
257	4	0.7	4	12	AAR11880	Fibrinogen recepto	430	3	0.7	4	15	AAR59892	2-amino-6,7-dihydr
258	4	0.7	4	12	AAR11140	Tetstsin peptide an	431	3	0.7	4	15	AAR55338	Inhibitor of tyros
259	4	0.7	4	12	AAR11235	Ketone analogue of	432	3	0.7	4	15	AAR55172	Sequence of sterol
260	4	0.7	4	12	AAR11236	Ketone analogue of	433	3	0.7	4	15	AAR64044	Acylated polyhydro
261	4	0.7	4	12	AAR12749	Antiallergic tetra	434	3	0.7	4	15	AAR57422	Ratpallin-3A fragm
262	4	0.7	4	12	AAR12868	Chromogenic substr	435	3	0.7	4	15	AAR60560	Vampire bat plasmi
263	4	0.7	4	12	AAR13803	Factor Xa substrat	436	3	0.7	4	15	AAR42559	Tri-/tetra-cyclic
264	4	0.7	4	12	AAR13306	Factor Xa substrat	437	3	0.7	4	15	AAR52983	Recom. EL receptor
265	4	0.7	4	12	AAR15768	Farnesyl protein t	438	3	0.7	4	15	AAR60813	Peptide fragment #
266	4	0.7	4	13	AAR22483	Neurotrophic facto	439	3	0.7	4	15	AAR69474	Integrin-inhibitin
267	4	0.7	4	13	AAR22486	Neurotrophic facto	440	3	0.7	4	15	AAR65922	Peptide used in th
268	4	0.7	4	13	AAR22487	Neurotrophic facto	441	3	0.7	4	15	AAR65923	Peptide used in th
269	4	0.7	4	13	AAR22499	Neurotrophic facto	442	3	0.7	4	15	AAR65924	Peptide used in th
270	4	0.7	4	13	AAR23394	Sequence of amphip	443	3	0.7	4	16	AAR93983	Anticopeptidic pep
271	4	0.7	4	13	AAR23259	Peptide for bindin	444	3	0.7	4	16	AAR93984	Anticopeptidic pep
272	4	0.7	4	13	AAR23683	Immunomodulator pe	445	3	0.7	4	16	AAR01707	Inhibitor of fibro
273	4	0.7	4	13	AAR23793	Selective factor 1	446	3	0.7	4	16	AAR74354	Human conserved BG
274	4	0.7	4	13	AAR25317	Cell contact inhib	447	3	0.7	4	16	AAR89858	Melanotrophic rele
275	4	0.7	4	13	AAR28903	Ratase partial amin	448	3	0.7	4	16	AAR37974	Cellular like neuro
276	4	0.7	4	13	AAR23270	Soybean glycinin d	449	3	0.7	4	16	AAR32148	Optical C-termina
277	4	0.7	4	14	AAR65685	Penicilliopepsin-in	450	3	0.7	4	16	AAR90881	Mu opioid receptor
278	4	0.7	4	14	AAR43677	Internalisation el	451	3	0.7	4	16	AAR87945	Beta amyloid (29-2
279	4	0.7	4	14	AAR37171	Aspartic proteinase	452	3	0.7	4	16	AAR80467	Artificial male st
280	4	0.7	4	14	AAR32380	Fibrinogen recepto	453	3	0.7	4	16	AAR80468	Artificial male st
281	4	0.7	4	14	AAR30163	MAB 13.1 variable	454	3	0.7	4	16	AAR90160	Neurotensin-relate
282	4	0.7	4	14	AAR22195	Proteolytic enzyme	455	3	0.7	4	16	AAR79523	Perfringolysin O C
283	4	0.7	4	14	AAR34100	Peptide which bind	456	3	0.7	4	16	AAR83195	Tetrapeptide bavin
284	4	0.7	4	14	AAR34038	Peptide which bind	457	3	0.7	4	16	AAR83201	Tetrapeptide bavin
285	4	0.7	4	14	AAR34090	Peptide which bind	458	3	0.7	4	16	AAR85668	Carbohydrate bindi
286	4	0.7	4	14	AAR34465	Bacillus lichenifo	459	3	0.7	4	16	AAR73037	Marine tyrosinase
287	4	0.7	4	14	AAR37365	Group 6 synthetic	460	3	0.7	4	16	AAR82911	Non-RGD, non-VISGR
288	4	0.7	4	14	AAR41972	Yeast endoplasmic	461	3	0.7	4	16	AAR82920	Non-RGD, non-VISGR
289	4	0.7	4	14	AAR38135	Protease non-labil	462	3	0.7	4	16	AAR72877	Yeast ppiase trypt
290	4	0.7	4	14	AAR38139	Protease non-labil	463	3	0.7	4	16	AAR72921	Yeast ppiase trypt
291	4	0.7	4	14	AAR38140	Protease non-labil	464	3	0.7	4	16	AAR65231	Fibrinogen/beta-ra
292	4	0.7	4	14	AAR38141	Protease non-labil	465	3	0.7	4	16	AAR71339	Transduction domain
293	4	0.7	4	14	AAR38127	Protease non-labil	466	3	0.7	4	16	AAR71343	IGEA motif of the
294	4	0.7	4	14	AAR38101	Protease labile fr	467	3	0.7	4	16	AAR67983	Thrombospondin-der
295	4	0.7	4	14	AAR38109	Protease non-labil	468	3	0.7	4	16	AAR79457	Mutant thrombospon
296	4	0.7	4	14	AAR38114	Protease non-labil	469	3	0.7	4	16	AAR69769	Thrombospondin pep
297	4	0.7	4	14	AAR38322	Diuretic hormone d	470	3	0.7	4	16	AAR66573	Anti-beta-endorphi
298	4	0.7	4	14	AAR38333	Diuretic hormone d	471	3	0.7	4	16	AAR66387	Hemolysin VPI BC
299	4	0.7	4	14	AAR38485	Human Phe de13219	472	3	0.7	4	16	AAP77829	Farnesyl transfera
300	4	0.7	4	14	AAR41956	Diuretic hormone d	473	3	0.7	4	16	AAP84228	Integrin-inhibitin
301	4	0.7	4	14	AAR43086	Diuretic hormone d	474	3	0.7	4	16	AAR83227	Integrin-inhibitin
302	4	0.7	4	14	AAR43091	Diuretic hormone d	475	3	0.7	4	16	AAR83227	Integrin-inhibitin
303	4	0.7	4	14	AAR43092	Diuretic hormone d	476	3	0.7	4	16	AAR87913	ATPase2AP peptide-1

500	4	19	AAW61337	PL1 binding transp	596	4	20	AAV28719	Human hepatitis C
501	4	19	AAW62050	Human erythropoiet	597	4	20	AAV28722	Human hepatitis C
502	4	19	AAW62052	Human erythropoiet	598	4	20	AAV31456	Anti-human consery
503	4	19	AAW61131	Peptide 10 used to	599	4	20	AAV31031	Carbohydrate antiq
504	4	19	AAW61136	Peptide 15 used to	600	4	20	AAV31114	Non-crosslinked pr
505	4	19	AAW61137	Peptide 16 used to	601	4	20	AAV31115	Non-crosslinked pr
506	4	19	AAW61143	Peptide 22 used to	602	4	20	AAV31051	Non-crosslinked pr
507	4	19	AAW61128	Peptide 7 used to	603	4	20	AAV30957	Non-crosslinked pr
508	4	19	AAW61147	Peptide 26 used to	604	4	20	AAV32466	Non-crosslinked pr
509	4	19	AAW51469	Pre sequence which	605	4	20	AAV25690	Membrane retaining
510	4	19	AAW51470	Pre sequence which	606	4	20	AAV25698	Membrane retaining
511	4	19	AAW51471	Pre sequence which	607	4	20	AAV24258	Somatostatin antag
512	4	19	AAW51472	Pre sequence which	608	4	20	AAV22418	Collagen I peptide
513	4	19	AAW51473	Pre sequence which	609	4	20	AAV27155	Peptide Seg ID No:
514	4	19	AAW51468	Pre sequence which	610	4	20	AAV23847	Single compound fo
515	4	19	AAW51458	Pre sequence which	611	4	20	AAV23612	Seq ID 103 of US9
516	4	19	AAW51459	Pre sequence which	612	4	20	AAV23461	V beta 6 clone fou
517	4	19	AAW51462	Pre sequence which	613	4	20	AAV23475	V beta 6 clone fou
518	4	19	AAW51463	Pre sequence which	614	4	20	AAV23267	Protein binding po
519	4	19	AAW51464	Pre sequence which	615	4	20	AAV17534	Human BRCAL intera
520	4	19	AAW51466	Pre sequence which	616	4	20	AAV17568	Hepatitis B virus
521	4	19	AAW52354	Truncated dncorphi	617	4	20	AAV17967	Hepatitis B virus
522	4	19	AAW52313	S. ariditaii strept	618	4	20	AAV21579	Integrin-binding p
523	4	19	AAW51184	Peptide chain of a	619	4	20	AAV17469	Antidepressant age
524	4	19	AAW51151	Cysteine containin	620	4	20	AAV09635	Antimicrobial cycl
525	4	19	AAW56816	Enzyme inhibitor p	621	4	20	AAV09630	Antimicrobial cycl
526	4	19	AAW51018	Hirudin-2, a thy	622	4	20	AAV15644	Peptide used to ma
527	4	19	AAW55758	Immunisation moti	623	4	20	AAV09606	Antimicrobial cycl
528	4	19	AAW55764	Immunisation moti	624	4	20	AAV09586	Antimicrobial cycl
529	4	19	AAW55746	Immunisation moti	625	4	20	AAV09592	Antimicrobial cycl
530	4	19	AAW55750	Immunisation moti	626	4	20	AAV09568	Antimicrobial cycl
531	4	19	AAW48196	Constatkin peptide	627	4	20	AAV09570	Antimicrobial cycl
532	4	19	AAW49978	Constatkin peptide	628	4	20	AAV09582	Antimicrobial cycl
533	4	19	AAW40438	Tobacco invertase	629	4	20	AAV09556	Antimicrobial cycl
534	4	19	AAW52428	Loop region used i	630	4	20	AAV14098	Affinity ligand fo
535	4	19	AAW52433	Loop region used i	631	4	20	AAV09403	Immunoreactive pepti
536	4	19	AAW52473	Beta turn region u	632	4	20	AAV09404	Immunoreactive pepti
537	4	19	AAW52385	Beta-turn region u	633	4	20	AAV09406	Immunoreactive pepti
538	4	19	AAW52389	Beta-turn region u	634	4	20	AAV09436	Endoplasmic reticu
539	4	19	AAW52495	Beta-turn region u	635	4	20	AAV09479	Blifobacterium bi
540	4	19	AAW52409	Beta-turn region u	636	4	20	AAV07508	RNA-34; cyclo(-L-G
541	4	19	AAW52350	Beta turn region u	637	4	20	AAV01252	CS5851813 peptide
542	4	19	AAW52371	Beta turn region u	638	4	20	AAV09144	Peptide #8 cleavaq
543	4	19	AAW53378	Rephata claspines r	639	4	20	AAW77455	Polylysine core of
544	4	19	AAW46026	Polylysine core ca	640	4	20	AAW78393	Synthetic Cys2-his
545	4	19	AAW47405	E. coli derived L-	641	4	20	AAW78376	Isoclectric focusi
546	4	19	AAW42372	Substrate #2 for i	642	4	20	AAW90256	RBC BIDS binding m
547	4	19	AAW45579	Peptide fragment o	643	4	20	AAW96758	Peptide sequence o
548	4	19	AAW43393	IAA wild type amir	644	4	20	AAW81200	Linker used in the
549	4	19	AAW46526	Peptide containing	645	4	20	AAW83543	Tumour necrosis fa
550	4	19	AAW46528	Peptide containing	646	4	20	AAW87608	Peptide derived fr
551	4	19	AAW46541	Peptide containing	647	4	20	AAW83484	Human growth hormo
552	4	19	AAW46542	Peptide containing	648	4	20	AAW83489	Human growth hormo
553	4	19	AAW46548	Peptide containing	649	4	20	AAW83469	Human growth hormo
554	4	20	AAV50606	Resin bound cyclic	650	4	20	AAW83453	Human growth hormo
555	4	20	AAV50607	Resin bound cyclic	651	4	20	AAW83466	Human growth hormo
556	4	20	AAV32118	Maize id gene null	652	4	20	AAW82596	Mouse Smad2 protei
557	4	20	AAV32118	Human IgG peptide	653	4	20	AAW81932	Elastin fragment u
558	4	20	AAV32428	Human melanoma an	654	4	20	AAH24354	Human prostate spe
559	4	20	AAV50349	Neutrophil-activat	655	4	20	AAH27267	Marine confluency
560	4	20	AAV50355	Neutrophil-activat	656	4	20	AAH27268	Marine confluency
561	4	20	AAV50357	Neutrophil-activat	657	4	20	AAH28273	Trypsin Modulating
562	4	20	AAV50364	Neutrophil-activat	658	4	20	AAH36145	Trypsin Modulating
563	4	20	AAV53991	HLA binding plu l	659	4	20	AAH36356	Wheat angiotensin-
564	4	20	AAV42566	collaspeptide compo	660	4	20	AAH24291	Prostate tumour as
565	4	20	AAV31858	Cell adhesion modu	661	4	20	AAH24295	Prostate tumour as
566	4	20	AAV39849	Elastin peptide fr	662	4	20	AAH24309	Prostate tumour as
567	4	20	AAV39841	Elastin peptide fr	663	4	20	AAH24317	Human H8087 casei
568	4	20	AAV40815	Amino acid loop 16	664	4	20	AAH28663	Chemotactic peptid
569	4	20	AAV30241	EGF/PLK 1 targetin	665	4	20	AAH28668	Chemotactic peptid
570	4	20	AAV30276	Icterin alpha-1 b	666	4	20	AAH29120	Cyclic human neutr
571	4	20	AAV31685	Icterin peptide used	667	4	20	AAH29123	Cyclic human neutr
572	4	20	AAV26241	Isolated peptide i	668	4	20	AAH29124	Cyclic human neutr

815	3	0.7	4	21	AAV64574	Desmoglein cell ad	898	3	0.7	4	22	AA885558	PTH2 receptor bind
816	3	0.7	4	21	AAV64544	Catharin related a	899	3	0.7	4	22	AA885566	Peptide used in me
817	3	0.7	4	21	AAV43661	Amino acid sequenc	899	3	0.7	4	22	AA885574	Peptide used in me
818	3	0.7	4	21	AAV43668	Amino acid sequenc	899	3	0.7	4	22	AA885579	Peptide used in me
819	3	0.7	4	22	AAV56072	Wilm's tumor (WIT	899	3	0.7	4	22	AA885591	HIV-1 Tat peptide
820	3	0.7	4	22	AAV56072	Melanoma antigen,	899	3	0.7	4	22	AA846645	Thrombospondin hep
821	3	0.7	4	22	AAV51890	ClpG-STH fusion pr	894	3	0.7	4	22	AA848337	Peptide spacer, S
822	3	0.7	4	22	AAV69157	Wheat protein deriv	895	3	0.7	4	22	AA865921	HF peptide fragmen
823	3	0.7	4	22	AAV68258	Wheat protein deriv	896	3	0.7	4	22	AA865945	HF peptide fragmen
824	3	0.7	4	22	AAV98726	Human peptide #201	897	3	0.7	4	22	AA866348	HF peptide fragmen
825	3	0.7	4	22	AAV51475	Anti-HIV peptide w	898	3	0.7	4	22	AA866349	HF peptide fragmen
826	3	0.7	4	22	AAV11869	Echinomys polialis	899	3	0.7	4	22	AA866959	Human peptide #7
827	3	0.7	4	22	AAV68961	Trifluoroacetic	900	3	0.7	4	22	AA867365	Human peptide #7
828	3	0.7	4	22	AAV79033	Amino acid sequenc	901	3	0.7	4	22	AA869143	M. catarrhalis 290
829	3	0.7	4	22	AAV96384	Catharin derived	902	3	0.7	4	22	AA830744	Peptide which is u
830	3	0.7	4	22	AAV64154	Hepatitis C virus	903	3	0.7	4	22	AA830745	Peptide which is u
831	3	0.7	4	22	AA884176	Peptide spacer lin	904	3	0.7	4	22	AA830754	Peptide which is u
832	3	0.7	4	22	AAV05115	Protein kinase C s	905	3	0.7	4	22	AA830760	Peptide which is u
833	3	0.7	4	22	AAV05194	Zinc finger protei	906	3	0.7	4	22	AA830762	Peptide which is u
834	3	0.7	4	22	AAV05426	Synthetic tetrapep	907	3	0.7	4	22	AA830764	Peptide which is u
835	3	0.7	4	22	AAV05436	Synthetic tetrapep	908	3	0.7	4	22	AA830765	Peptide which is u
836	3	0.7	4	22	AAV05628	Synthetic glycosyl	909	3	0.7	4	22	AA830773	Peptide which is u
837	3	0.7	4	22	AAV07257	Peptide #22 used i	910	3	0.7	4	22	AA830775	Peptide which is u
838	3	0.7	4	22	AAV07260	Peptide #25 used i	911	3	0.7	4	22	AA830776	Peptide which is u
839	3	0.7	4	22	AAV07269	Peptide #34 used i	912	3	0.7	4	22	AA830791	Peptide which is u
840	3	0.7	4	22	AAV99864	BRA binding domain	913	3	0.7	4	22	AA851332	Angiotensin-conver
841	3	0.7	4	22	AAV95441	C-terminal sequenc	914	3	0.7	4	22	AA851333	Angiotensin-conver
842	3	0.7	4	22	AAV85405	Peptide linker for	915	3	0.7	4	22	AA860177	P. falciparum circ
843	3	0.7	4	22	AAV51682	Proactive tetrapep	916	3	0.7	4	22	AA876444	BBB2 binding pept
844	3	0.7	4	22	AAV04662	Oploid receptor sp	917	3	0.7	4	22	AA850875	Integrin recogniti
845	3	0.7	4	22	AAV04665	Oploid receptor sp	918	3	0.7	4	22	AA851213	Caspase cleavage s
846	3	0.7	4	22	AAV99514	Angiotensin I conv	919	3	0.7	4	22	AA851323	Caspase cleavage s
847	3	0.7	4	22	AAV05141	Marine W-CSP prot	920	3	0.7	4	22	AAV59134	Peptide #2 recogni
848	3	0.7	4	22	AAV98389	Buffalo rat liver	921	3	0.7	4	22	AAV59314	Bov colostrinin pe
849	3	0.7	4	22	AAV98996	Yeast C-terminal E	922	3	0.7	4	22	AAV59371	Elastin repeating
850	3	0.7	4	22	AAV63425	Casain related pep	923	3	0.7	4	22	AAV64022	Peptide 3 for pept
851	3	0.7	4	22	AAV63427	Casain-related pep	924	3	0.7	4	22	AAV15587	Oligomeric compou
852	3	0.7	4	22	AAV97411	Beta-glucosyl epito	925	3	0.7	4	22	AAV48440	Fluorescent electr
853	3	0.7	4	22	AAV97441	Collagen IV alpha	926	3	0.7	4	22	AAV48759	Fluorescent electr
854	3	0.7	4	22	AAV94942	Peptide based vac	927	3	0.7	4	22	AAV48762	Fluorescent electr
855	3	0.7	4	22	AAV70927	Melanocortin recep	928	3	0.7	4	22	AAV49332	Proteasome-targeti
856	3	0.7	4	22	AAV70946	Melanocortin recep	929	3	0.7	4	22	AAV55550	T cell surface rec
857	3	0.7	4	22	AAV74957	Octopus minor brai	930	3	0.7	4	22	AAV55555	T cell surface rec
858	3	0.7	4	22	AAV96974	Mammalian stem cel	931	3	0.7	4	22	AAV55610	Monocyte surface r
859	3	0.7	4	22	AAV62198	Human gene 10-enco	932	3	0.7	4	22	AAV55645	Endothelial cell s
860	3	0.7	4	22	AAV67729	Amino acid sequenc	933	3	0.7	4	22	AAV55680	Endothelial cell s
861	3	0.7	4	22	AAV81104	CAC25 inhibitory i	934	3	0.7	4	22	AAV55685	Endothelial cell s
862	3	0.7	4	22	AAV91532	Oploid peptide SEQ	935	3	0.7	4	22	AAV55597	CD49f C-terminal c
863	3	0.7	4	22	AAV91462	Oploid peptide SEQ	936	3	0.7	4	22	AAV55402	CD49f C-terminal c
864	3	0.7	4	22	AAV91741	Oploid peptide SEQ	937	3	0.7	4	22	AAV57457	LPAP C-terminal co
865	3	0.7	4	22	AAV92182	Laminin fragment S	938	3	0.7	4	22	AAV57432	CD148 C-terminal c
866	3	0.7	4	22	AAV92399	Mucopolysacchar pep	939	3	0.7	4	22	AAV57527	CD66b and CD66c C-
867	3	0.7	4	22	AAV92490	Mucopolysacchar pep	940	3	0.7	4	22	AAV57532	CD195 C-terminal c
868	3	0.7	4	22	AAV67710	Amino acid sequenc	941	3	0.7	4	22	AAV57807	CD49f C-terminal c
869	3	0.7	4	22	AAV84314	Vasculostatin pep	942	3	0.7	4	22	AAV57812	CD49f C-terminal c
870	3	0.7	4	22	AAV84384	Synthetic peptide,	943	3	0.7	4	22	AAV57867	LPAP C-terminal co
871	3	0.7	4	22	AAV73920	Peptide SEQ IV NC	944	3	0.7	4	22	AAV57902	CD148 C-terminal c
872	3	0.7	4	22	AAV03020	Peptide #3 insert	945	3	0.7	4	22	AAV57937	CD66b and CD66c C-
873	3	0.7	4	22	AAV72355	Colostrinin deriv	946	3	0.7	4	22	AAV57942	CD195 C-terminal c
874	3	0.7	4	22	AAV73419	Human PC-LENTIN R	947	3	0.7	4	22	AAV48424	Human ICAM-1 pepti
875	3	0.7	4	22	AAV87340	Breast-cancer asso	948	3	0.7	4	22	AAV48425	Human ICAM-1 pepti
876	3	0.7	4	22	AAV87340	Breast-cancer asso	949	3	0.7	4	22	AAV328247	P2X2 peptide frag
877	3	0.7	4	22	AAV69541	Peptide linker, c	950	3	0.7	4	22	AAV29559	Adrenic modulator
878	3	0.7	4	22	AAV69985	Stabilising peptid	951	3	0.7	4	22	AAV82516	Mouse orphan recep
879	3	0.7	4	22	AAV70327	Osteoclastogenesis	952	3	0.7	4	22	AAV86851	Cathepsin-associat
880	3	0.7	4	22	AAV73599	Colostrinin peptid	953	3	0.7	4	22	AAV86867	Transport molecule
881	3	0.7	4	22	AAV73541	Colostrinin peptid	954	3	0.7	4	22	AAV86873	Transport molecule
882	3	0.7	4	22	AAV73527	Mammalian interfer	955	3	0.7	4	22	AAV86876	Transport molecule
883	3	0.7	4	22	AAV73529	Mammalian interfer	956	3	0.7	4	22	AAV86890	Transport molecule
884	3	0.7	4	22	AAV72701	Repetitive protein	957	3	0.7	4	22	AAV86917	Transport molecule
885	3	0.7	4	22	AAV72762	Repetitive protein	958	3	0.7	4	22	AAV355101	Anti IL-19 antibod
886	3	0.7	4	22	AAV80155	PTH2 receptor bind	959	3	0.7	4	22	AAV65421	Representative cla
887	3	0.7	4	22	AAV80192	PTH2 receptor bind	960	3	0.7	4	22	AAV10145	Sequence of pentap

[illegible]

2000

[illegible]

PA (RES) UNIV CALIFORNIA.
 XX
 PT Petersen C, Huang J;
 XX
 DR WPI: 2001 38923/11.
 XX
 PT New vaccines and cryptosporidium protein for diagnosis and treatment of
 PT cryptosporidium species infection .
 XX
 PS Example 2: Column 19: 32pp; English.
 XX
 CC The invention relates to vaccines and cryptopain protein for diagnosis
 CC and treatment of cryptosporidium species infection. cryptopain protein
 CC can be expressed by standard recombinant methodology. The vaccines
 CC comprising the cryptopain protein or its fragments are useful for active
 CC immunization of animals and humans against cryptosporidium infection, or
 CC for production of passive immune products in admixture with an adjuvant.
 CC cryptopain can be used for prophylactic, therapeutic, diagnostic and
 CC detection purposes. The present sequence represents a peptide fragment
 CC of the C. parvum cryptopain protein.
 XX
 SQ Sequence 5 AA:
 Query Match 1.2%, Score 5, DB 22, Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0.
 QY 360 VNNSW 364
 DB 1 VNNSW 5
 RESULT 3
 AAP90988
 ID AAP90988 standard; protein: 6 AA.
 AC AAP90988;
 XX
 DI 06-JUN 1990 (first entry)
 XX
 DE Binding receptor with selectivity for a target ligand, borne by
 DE an article for inactivating toxic materials.
 XX
 KW Toxic material, inactivator, organophosphorous spds., nerve poison,
 KW pesticide; decontaminant; military.
 XX
 PN WG902920-A
 XX
 PD 06-APR-1989.
 XX
 PF 04-OCT-1988; 98W0-0503422.
 XX
 PR 05-OCT-1987; 8706-0106912
 XX
 PA (LITT) LITTLE AB INC.
 XX
 PI Taylor RF;
 XX
 DR WPI: 1989-114395/15.
 XX
 PT Article for inactivating toxic materials, eg organophosphorous spds., -
 PT comprises solid carrier bearing target ligand binding receptor
 PI and ligand-degrading receptor, pref. enzyme
 XX
 PS Claim 7g; page 42; 57pp; English.
 XX
 CC The article for inactivating a toxic material comprises a solid carrier
 CC bearing a first receptor which binds the target ligand and a second
 CC receptor which degrades the target ligand. This synthetic peptide is
 CC a preferred first receptor. The article may be used for covering
 CC surfaces to protect or decontaminate the surface. The article is esp. for
 CC degrading toxic organic spds., esp. organophosphorous spds., (eg
 CC pesticides and nerve poisons, bacteria and viruses, in environmental,

CC chemical, military and industrial settings.
 XX
 SQ Sequence 6 AA:
 Query Match 1.2%, Score 5, DB 10, Length 6;
 Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0.
 QY 197 CGSCW 201
 DB 2 CGSCW 6
 RESULT 4
 AAK10048
 ID AAK10048 standard, pep-idb: 6 AA.
 AC AAK10048;
 XX
 DT 15-MAR 1991 (first entry)
 XX
 DE Hypotensive oligopeptide.
 XX
 FW 2-ai Hypotensive; fig.
 XX
 OS Ficus carica.
 XX
 PN JF02282394-A.
 XX
 PD 19-NOV-1990.
 XX
 PF 24-APR-1989; 89JP-0104265.
 XX
 PR 24-APR-1989; 89JP-0104265.
 XX
 PA (AGEN) AGENCY OF IND SCI TECH.
 PA (SHOS) SHOWA SANGYO KK.
 XX
 DR WPI: 1991 004480/01.
 XX
 PI New oligopeptide hypotensive diag - based on alanine, valine,
 PI asparagine, proline, isoleucine and arginine
 XX
 TS Claim 1, Page 843, 9pp, Japanese.
 XX
 CC Peptides may be derived from extract of fig. and are hypotensive
 CC agents.
 XX
 SQ Sequence 6 AA:
 Query Match 1.2%, Score 5, DB 12, Length 6;
 Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0.
 QY 188 VNPIR 192
 DB 2 VNPIR 6
 RESULT 5
 AAR57003
 ID AAR57003 standard, peptide, 6 AA.
 AC AAR57003;
 XX
 DT 15-FEB-1995 (first entry)
 XX
 DE N-terminal fragment of human histo blood group A transferase.
 XX
 KW Blood, group, determinant, antigen, erythrocyte, oligosaccharide;
 KW glycosylate, glycosphingolipid, glycoprotein, glycosyltransferase;
 KW transferase.
 XX

Query Match 1.00%; Score 4; DB 14; Length 4;
 Best Local Similarity 100.00%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 3883 004
 II I I
 Db 1 0000 4

RESULT 10

AAK67642
 ID AAK67642 standard; Protein; 4 AA.

XX AAK67642
 AC AAK67642

XX 27-SEP-1994 (first entry)

XX Sequence of single-chain antigen-binding protein.

XX Antibody; multivalent; variable region; heavy chain; light chain;

XX linker; hinge.

XX Synthetic.

XX Key Location/Qualifiers

PI Misc-difference 2 /label residue 218

XX W09411161 A.

XX 11-JUN-1994

XX 20-NOV-1992; 9280-0539995.

XX 25-NOV-1991; 9108-0796946.

XX (PDB) ENZYME INC.

XX Bird RE; Filpula D; Hardman K; Kollence M; Whitlow MD;

XX Wood JE;

XX WPI; 1994-19699/24.

XX New multivalent antigen-binding protein e.g. contg. CC49 Mab
 PI regions useful in diagnosis; for destroying blood clots and
 PI targeting cytotoxic agents or enzymes to tumor cells

XX Example; Table 5, page 46; 118pp; English.

XX Free cysteines were engineered into the C-terminal of the
 CC 4-4-20/212 single-chain antigen-binding protein, in order to
 CC chemically crosslink the protein. The design was based on the
 CC hinge region found in antibodies between the CH1 and CH2 regions.
 CC In order to try to restore antigenicity in humans, the hinge
 CC sequence of the most common IgG class, IgG1, was chosen. In the
 CC design for the hinge region, the C-terminal serine in the
 CC 4-4-20/212 single-chain antigen-binding protein was made the first
 CC serine of the hinge and the second residue of the hinge was changed
 CC from a cysteine to a serine. The 4-4-20 Fab structure was examined
 CC and it was determined that the C-terminal sequence
 CC GluH216-ProH217-ArgH218, was part of the CH1 region and that the
 CC hinge between CH1 and CH2 starts with ArgH218 or GlyH219 in the
 CC mouse 4-4-20 14G2A antibody.

XX Sequence 4 AA;

Query Match

Best Local Similarity 1.00%; Score 4; DB 14; Length 4;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VIVS 55
 II I I I

Db 1 VIVS 4

RESULT 9

AAK46063
 ID AAK46063 standard; peptide; 4 AA.

XX AAK46063

XX 22-JUL-1994 (first entry)

XX EHL' derived peptide CNBr-4.

XX reduced; derivative; deriv; EHL; too in; larvicidal; insecticidal;
 XX larvicide; insecticide; treatment; prevention; viral; infection.

XX Branthis hyemalis.

XX W09402513-A.

XX 04-FEB-1994.

XX 22-JUL-1993; 93W0-US06945.

XX 24-JUL-1992; 92US-0919464.

XX (PDB) FINDER HI-BRED INT INC.

XX Kumar MA; Rao AG;

XX WPI; 1994-048747/06.

XX New reduced derivative of Branthis hyemalis lectin having larvicidal
 XX and insecticidal activity and also used to treat or prevent viral
 XX infection, partic. in plants

XX Example 10; Page 18; 30pp; English.

XX The sequence is that of a cyanogen bromide peptide digest of a

CC completely reduced deriv. of Branthis hyemalis lectin (EHL), EHL'.
 CC EHL' epds. retain the larvicidal and insecticidal activity of EHL.
 CC but lack the protein synthesis of the parent epd. The epds. can also
 CC be used for treating or preventing viral infection, partic. in plants.

XX Sequence 4 AA;

Query Match 1.00%; Score 4; DB 15; Length 4;

Best Local Similarity 100.00%; Pred. No. 6.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 DEDT 453

II I I I I

Db 1 DEDT 4

RESULT 10

AAK16437
 ID AAK16437 standard; peptide; 4 AA.

XX AAK16437

XX 06-AUG-1999 (first entry)

XX Synthetic cyclosporin A (CSA) peptide analogue.

XX Immunostimulatory peptide analogue; cyclosporin A; CSA;

XX immunosuppressive; immune cell stimulation; lymphocyte proliferation;

XX Lymphocyte activity; HIV infection.

XX Synthetic.

XX W09606857-A1.

XX

Query Match 1.0%, Score 4, DB 18, Length 4;
 Best Local Similarity 100.0%, Pred. No. 6.4e+05;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0.

QY 155 QWER 158
 DB 1 DDER 4

RESULT 13
 AAY42618
 ID AAY42618 standard; peptide; 4 AA.
 XX
 AC AAY42618;
 XX
 DI 10-JAN-2000 (first entry)
 XX
 DE Human IgE variant fragment
 XX
 KW Immunoglobulin E; IgE; antagonist; FcεpsilonRI receptor; human; bds;
 KW Receptor-binding; binding determinant sequence; anti-IgE antibody;
 KW allergic disease.
 XX
 OS Homo sapiens.
 XX
 FN US5945709 A.
 XX
 FE 12-OCT-1999.
 XX
 FE 21-APR-1994; 94US-0722530
 XX
 FE 14-AUG-1991; 91US-0744768.
 XX
 FE 07-JAN-1994; 94US-0178583.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Jardieu PM, Presta IG;
 XX
 FE WPI, 1999-576041/49

Immunoglobulin E variants as peptide antagonists useful for raising and screening anti-immunoglobulin E (IgE) antibodies, in the isolation and purification of FcεpsilonRI receptor and in the treatment of allergic diseases.

Example 1, Column 50, 57pp, English.

The invention relates to immunoglobulin E (IgE) antagonists comprising one or more of the FcεpsilonRI receptor binding determinant sites of human IgE. The antagonists include IgE variants comprising an immunoglobulin E template and binding determinant sequences (bds) CDbs, ERbds and the sequence shown in AAY42581. The CDbs (CD loop binding determinant sequence) are selected from the sequences shown in AAY42577 and the ERbds (EF loop binding determinant sequence) are selected from the sequences shown in AAY42578 Y42580. The variants are useful in raising and screening anti-IgE antibodies, in the isolation and purification of FcεpsilonRI receptor and in the treatment and prophylaxis of allergic diseases.

XX
 SO Sequence 4 AA;

Query Match 1.0%, Score 4, DB 20, Length 4;
 Best Local Similarity 100.0%, Pred. No. 6.4e+05;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0.

QY 207 AALE 210
 DB 1 AALE 4

RESULT 14
 AAY29992

ID AAW88092 standard; peptide; 4 AA.
 XX
 AC AAW88092;
 XX
 DI 13-APR-1999 (first entry)
 XX
 DE C-terminal sequence of an antigen binding protein.
 XX
 KW Antigen binding protein; variable region; hinge region;
 KW antibody light chain; antibody heavy chain; medical diagnosis;
 KW organ imaging; immunotherapy; targeting; immunoassay;
 KW immunoaffinity purification.
 XX
 OS Unidentified.
 XX
 PN US5869620-A.
 XX
 PD 09-FEB-1999.
 XX
 FE 22-FEB-1995; 95US-0392338.
 XX
 FE 20 NOV 1992; 92US 0989846.
 XX
 FE 02-SEP-1986; 86US-0902971.
 XX
 FE 02-SEP-1987; 87US-0092110.
 XX
 FE 19-JAN-1989; 89US-0299617.
 XX
 FE 25-APR-1990; 90US-0512910.
 XX
 FE 25 NOV 1991; 91US 0796926.
 XX
 FE 22-FEB-1995; 95US-0392338.
 XX

(ENZO-) ENZON INC.
 XX
 PI Bird RE, Filipula D, Hardman KD, Kollence M, Whitlow MD;
 PI Wood JF;
 XX
 FE WPI, 1999-152879/13.
 XX
 FE Multivalent antigen-binding proteins - comprising two or more
 PT single-chain Fv molecules
 XX
 FE Example 5, Column 21; 71pp, English.

The present sequence represents the C-terminal sequence of an antigen binding protein of the invention. The antigen binding protein is in a multivalent form comprising two or more single-chain molecules, each comprising a polypeptide with the antigen binding portion of the variable region of an antibody light chain and a heavy chain, where the polypeptides are linked together by a peptide linker. The antigen-binding protein is used in compositions that are useful for medical diagnosis, organ imaging, immunotherapy, targeting therapeutic or diagnostic agents, immunoassays and immunoaffinity purification.

XX
 SO Sequence 4 AA;
 Query Match 1.0%, Score 4, DB 20; Length 4;
 Best Local Similarity 100.0%, Pred. No. 6.4e+05;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VTVS 55
 DB 1 VTVS 4

RESULT 15
 AAB15757
 ID AAB15757 standard; peptide; 4 AA.
 XX
 AC AAB15757;
 XX
 DI 02-FEB-2001 (first entry)
 XX
 DE Nuclear translocation signal sequence.
 XX
 FE Targeted drug delivery; PAM, biological molecule; ophthalmological;

PA (CELL-) CELLOMICS INC.
 XX Giuliano KA, Kapur R;
 XX WPI: 2000-594086/56.
 DR N-PSDB: AAA93386.
 XX
 PT Automated cell-based characterization of toxin by contacting cells
 PT containing luminescent reporter molecules with test substance and
 PT analyzing optically -
 XX
 XX Example 11: Fig 24a; 24pp; English.
 XX
 XX The invention relates to systems, methods and reagents for cell-based
 XX screening or detection of compounds which affect particular biological
 XX functions. The methods of the invention utilize fluorescent bioreceptor
 XX molecules which, when acted on by a compound of interest, cause an
 XX alteration in the cellular distribution of at least the fluorescent
 XX moiety. In one embodiment, the biosensors comprise heat shock proteins
 XX (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent
 XX protein (GFP), or derivatives thereof). Such biosensors are located in
 XX the cytoplasm, but on stress activation, translocate to the nucleus. In
 XX another embodiment, bioreceptor proteins can be used to detect protease
 XX activity. Such protease-inducible fusion proteins comprise one or more
 XX fluorescent proteins, a recognition signal which is cleaved by the
 XX protease, and at least one cellular localisation signal. The latter two
 XX components may be components of a single protein which is acted upon by
 XX the protease, or may be from heterologous sources. Due to the
 XX localisation signal, the bioreceptor protein is localised to a
 XX particular region of the cell. Once acted on by the protease of interest,
 XX the fluorescent protein is released from the localised location, and
 XX is free to migrate to other locations within the cell. The presence of a
 XX second localisation signal attached to the fluorescent protein enables
 XX the fluorescent protein to be directed to a different cellular
 XX compartment after cleavage of the protease recognition sequence. The
 XX change in distribution of the fluorescent protein can be detected using
 XX imaging methods with a high degree of spatial resolution. The methods
 XX and biosensors of the invention can be used to investigate a wide range
 XX of cellular activities and to screen compounds which modulate these
 XX activities. Biosensors containing a recognition site for caspase, for
 XX example, may be used for the screening of compounds which modulate
 XX apoptosis, while biosensors containing other protease recognition sites
 XX may be used for the detection of proteolytic toxins (such as anthrax
 XX lethal factor). The method provides improved target validation and
 XX candidate compound optimisation by combining many cell screening formats
 XX with fluorescence-based molecular reagents and computer-based feature
 XX extraction, data analysis and automation, resulting in increased
 XX quantity and speed of data collection and faster evaluation of drug
 XX candidates. Sequences AAH2886-607900 and AAP2225 represent protease
 XX recognition sites which may be used as components of biosensor fusion
 XX proteins of the invention.
 XX
 XX Sequence 4 AA:
 XX
 XX Query Match 1.0%; Score 4; DB 21; Length 4;
 XX Best local Similarity 100.0%; Pred. No. 6.4e+05;
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 XX
 XX Q7 315 IQNS 218
 XX IIII
 XX Db 1 IQNS 4
 XX
 XX RESULT 18
 XX AAB07928
 XX ID AAB07928 standard; Peptide: 1 AA
 XX AC AAB07928;
 XX
 XX 14-NOV-2000 (first entry)
 XX
 XX C-terminal of single-chain antigen binding protein
 XX

KW Single-chain antigen-binding protein; C049 monoclonal antibody;
 KW immunosassay; protein purification; biosensor; IgG2A, ss.
 XX
 XX Synthetic.
 XX
 XX OS6103889-A.
 XX
 XX 15-AUG-2000.
 XX
 XX 14-OCT-1998; 98US-0172019.
 XX
 XX 20-NOV-1992; 92US-099846.
 XX
 XX 20-SEP-1995; 95US-029338.
 XX
 XX 25-NOV-1991; 91US-0796936.
 XX
 XX (EMCO-) ERCON INC.
 XX
 XX WHITLOW MD, BIRD RE, FILPULA D, HARDMAN KD;
 XX
 XX WPI: 2000-548439/59.
 XX
 XX Reel molecule add molecule encoding single chain antigen-binding
 XX proteins; multivalent antigen-binding proteins comprising the
 XX single chain proteins useful as the specificity to diagnostically
 XX effective agents -
 XX
 XX Example 5; Coloma 22; 23pp; English.
 XX
 XX The specification describes a single-chain antigen-binding protein,
 XX which comprises polypeptides having the antigen binding portions of the
 XX variable region of a heavy or light chain of a 67% monoclonal antibody
 XX having A07 deposit No. HA-9479 and a peptide linker linking the
 XX polypeptides into a single chain protein. The multivalent antigen binding
 XX proteins are useful for diagnosis, treatment of a medical condition, as
 XX a carrier to image the specific bodily organs of an animal, for
 XX immunoassays and in purification and biosensors. The present sequence
 XX represents the 5' terminal of a single chain antigen-binding protein
 XX of the invention.
 XX
 XX Sequence 4 AA:
 XX
 XX Query Match 1.0%; Score 4; DB 21; Length 4;
 XX Best local Similarity 100.0%; Pred. No. 6.4e+05;
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 XX
 XX Q7 52 VTWS 55
 XX IIII
 XX Db 1 VTWS 4
 XX
 XX RESULT 19
 XX AAY79597
 XX ID AAY79597 standard; Peptide: 4 AA.
 XX
 XX AC AAY79597;
 XX
 XX 14-AUG-2000 (first entry)
 XX
 XX Protease 7 substrate recognition sequence.
 XX
 XX
 XX Protease; biosensor; caspase 7; substrate recognition sequence;
 XX cell screening; assay; analysis; drug discovery.
 XX
 XX Unidentified.
 XX
 XX WC200006408-A2.
 XX
 XX 11-MAY-2000.
 XX
 XX 29-OCT-1999; 99WO-0525431.
 XX
 XX 30-OCT-1998; 98US-0106308.
 XX
 XX 26-MAY-1999; 99US-0146078.

00 The invention relates to the production of multivalent antigen binding
01 proteins from polypeptides comprising the antigen-binding portions of
02 antibody heavy and light chain variable regions (VH and VL). A single
03 chain molecule comprises VH and VL regions which are joined by a linker
04 peptide. In the method of the invention, the single chain
05 antigen-binding molecules are dissociated and then reassociated; the
06 multivalent antigen-binding proteins are separated from the single-chain
07 molecules and recovered. The multivalent antigen-binding proteins
08 produced can be used to diagnose a medical condition, as a carrier to
09 image the specific bodily organs of an animal, to treat a medical
10 condition and if radiolabelled in immunoassays and immunoaffinity
11 purifications. The multivalent proteins have enhanced binding ability as
12 their multivalent form provides more binding sites per molecule and they
13 can be used as multispecific binding molecules. Antigen-binding proteins
14 are smaller than whole antibodies so enhanced clearing from serum is
15 achieved which may provide lower background in imaging applications and
16 they may be less immunogenic. Multivalent antigen-binding proteins may
17 penetrate tumours better than monoclonal antibodies giving better tumour
18 fighting ability. The multivalent proteins lack the Fc component present
19 in antibodies which has binding sites for cells which include liver and
20 spleen cells and the absence of this component may reduce accumulation
21 of the protein in non-target tissues. The proteins are easily produced
22 and engineered compared with the Kohler and Milstein method used to
23 produce whole antibodies. Sequences AAY8315-Y8319 represent antibody
24 hinge regions. Sequences AAY80915-Y80916 represent the amino acid
25 sequences of hinge regions from murine and human IgGs (immunoglobulin G).
26 The human full sequence (AAY80916) was used as the basis for the design
27 of hinge regions (AAY80917-Y80919) for use in the construction of
28 multivalent antigen-binding proteins.

XX Sequence 4 AA;

Query Match 1.0%; Score 4; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VIVS 55
III
DB 1 VIVS 4

RESULT 22

AAE09724
ID AAY98724 standard; Peptide: 4 AA.

AC AAE09724;

AD AAY98724;

DI 24-JAN-2002 (first entry)

DE Human peptide #1999 exceeded by a SNP oligonucleotide.

EE Immunosuppressive; immunomodulatory; anti-inflammatory; cytostatic;
KW neuroprotection; antimicrobial; gene therapy; vasculase; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinase; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease

FS Homo sapiens.

IN W20014744 A2.

IL 16-JUL-2001.

IF 28-DEC-2000; 2000W0-US35498.

XX 28-DEC-1999; 9908-0173419.

PR 27-DEC-2000; 2000B3-0173419.

XX (CURA-) CURAGEN CORP.

PA

RI Shimkets RA, Leach M;
XX WI, 2001 455216, 20.
IR
XX Polymorphic nucleic acids encoding e.g. amylase, cyclin, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX Disclosure: Page 4106; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
00 encoding polymorphic variants of proteins related to amylases, amyloid
01 proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
02 polymerase, oncogenes, histones, kinases, colony stimulating factors,
03 complement related proteins, cytochromes, kinases, cytokines,
04 interleukins, G protein coupled receptors and thioesterases.
05 The present sequence is a peptide encoded by one such oligonucleotide.
06 The oligonucleotides and the peptides encoded by them may be used in the
07 prevention, diagnosis and treatment of diseases associated with
08 inappropriate expression of the proteins listed above. Disorders that may
09 be prevented, diagnosed and/or treated include multi-factorial diseases
10 with a genetic component, such as autoimmune diseases (e.g. rheumatoid
11 arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
12 and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
13 brain, breast, colon and kidney, leukaemia), diseases of the nervous
14 system and an infection of pathogenic organisms.

XX Sequence 4 AA;

Query Match 1.0%; Score 4; DB 22; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KSSR 164
IIII
DB 1 KSSR 4

RESULT 23

AAE09476
ID AAE05476 standard; peptide: 4 AA.

AC AAE05476;

AD 24-SPP-2001 (first entry)

DE Synthetic mini (connecting) C-peptide #3.

EE Insulin precursor; IP: connecting peptide; C-peptide; A chain;
KW B chain; increased production; folding stability; therapy;
KW diabetes mellitus.

CS Synthetic.

XX W20014742-A1.

PR 12-JUL-2001.

XX 04-DEC-2000; 2000W0-DK00665.

XX 29-DEC-1999; 9908-0001868.

PR 17-MAR-2000; 2000DK-0000440.

XX (NOVO) NOVO NORDISK AS.

XX Kjelsgaard TB, Ludvigsen S, Kaarsholm NC;

XX WI, 2001-451787/48.

XX Insulin precursor analogs with connecting peptides comprising aromatic
PT amino acids, and which have increased folding stability, useful for
PT treating diabetes -


```

11 22-DEC-1994 (first entry)
12 Cyclic-(D-Val-L-Phe-L-Asp-D-Ala-L-Arg).
13 cyclopeptide; thrombosis; heart attack; arteriosclerosis;
14 inflammation; apoplexy; angina pectoris; tumour; therapy;
15 metabolic disease; cyclic.
16 Synthetic.
17
18 Key Location/Qualifiers
19 Modified-site 1..5
20 /note= "the amino group of L-Val(1) forms a peptide
21 bond with the carboxy group of L-Arg(5)"
22
23 Misc-difference 1
24 /note= "D-form residue"
25
26 Misc-difference 4
27 /note= "D-form residue"
28
29 EP596350-A.
30
31 11-MAY-1994.
32
33 23-OCT-1993. 93EP-0117204.
34
35 06-NOV-1992. 92DE-4237456.
36 (MERE ) MERCK PATENT GMBH.
37
38 Diotenbach B., Felding-Habermann B., Haubner K., Hoelzemann G.;
39 Joneczyk A., Kessler H., Melzer G., Eppmann F., Wermuth J.;
40 WPI: 1994 152792/19.
41
42 New cyclopeptide(s) of five aminoacid(s) - are useful in
43 pharmaceutical preparations and in the treatment of metabolic
44 diseases, thrombosis and heart attacks
45
46 Example 1: Page 6; 7pp; German.
47
48 This peptide is a specific example of a generic cyclic pentapeptide
49 formula. The cyclopeptides are useful for the prophylaxis and
50 treatment of metabolic diseases, thrombosis, heart attacks,
51 arteriosclerosis, inflammation, angina pectoris and tumours.
52 See AAR52716 R52720 for specifically claimed examples.
53
54 Sequence 5 AA:
55
56 Query Match 1.0%; Score 4; DB 15; Length 5;
57 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
58 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
59
60 329 VFDA 332
61 1111
62 1 VFDA 4
63
64 RESULT 32
65 AAR52730
66 ID AAR52730 standard; peptide; 5 AA.
67 AC AAR52730;
68
69 22-DEC-1994 (first entry)
70
71 Cyclic-(L-Val-L-Phe-L-Asp-D-Ala-L-Arg).
72
73 cyclopeptide; thrombosis; heart attack; arteriosclerosis;
74 inflammation; apoplexy; angina pectoris; tumour; therapy;
75 metabolic disease; cyclic.
76 Synthetic.
77
78 Key Location/Qualifiers
79 Modified-site 1..5
80 /note= "the amino group of L-Val(1) forms a peptide
81 bond with the carboxy group of D-Arg(5)"
82
83 Misc-difference 4
84 /note= "D form residue"
85
86 Misc-difference 5
87 /note= "D-form residue"
88
89 EP596350-A.
90
91 11-MAY-1994.
92
93 23-OCT-1993. 93EP-0117204.
94
95 06-NOV-1992. 92DE-4237456.
96 (MERE ) MERCK PATENT GMBH.
97
98 Diotenbach B., Felding-Habermann B., Haubner K., Hoelzemann G.;
99 Joneczyk A., Kessler H., Melzer G., Eppmann F., Wermuth J.;
100 WPI: 1994 152792/19.
101
102 New cyclopeptide(s) of five aminoacid(s) - are useful in
103 pharmaceutical preparations and in the treatment of metabolic
104 diseases, thrombosis and heart attacks
105
106 Example 1: Page 6; 7pp; German.
107
108 This peptide is a specific example of a generic cyclic pentapeptide
109 formula. The cyclopeptides are useful for the prophylaxis and
110 treatment of metabolic diseases, thrombosis, heart attacks,
111 arteriosclerosis, inflammation, angina pectoris and tumours.
112 See AAR52716 R52720 for specifically claimed examples.
113
114 Sequence 5 AA:
115
116 Query Match 1.0%; Score 4; DB 15; Length 5;
117 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
118 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
119
120 329 VFDA 332
121 1111
122 1 VFDA 4
123
124 RESULT 32
125 AAR52730
126 ID AAR52730 standard; peptide; 5 AA.
127 AC AAR52730;
128
129 22-DEC-1994 (first entry)
130
131 Cyclic-(L-Val-L-Phe-L-Asp-D-Ala-L-Arg).
132
133 cyclopeptide; thrombosis; heart attack; arteriosclerosis;
134 inflammation; apoplexy; angina pectoris; tumour; therapy;
135 metabolic disease; cyclic.
136 Synthetic.
137
138 Key Location/Qualifiers
139 Modified-site 1..5
140 /note= "the amino group of L-Val(1) forms a peptide
141 bond with the carboxy group of D-Arg(5)"
142
143 Misc-difference 4
144 /note= "D form residue"
145
146 Misc-difference 5
147 /note= "D-form residue"
148
149 EP596350-A.

```

```

11 22-DEC-1994 (first entry)
12 Cyclic-(D-Val-L-Phe-L-Asp-D-Ala-L-Arg).
13 cyclopeptide; thrombosis; heart attack; arteriosclerosis;
14 inflammation; apoplexy; angina pectoris; tumour; therapy;
15 metabolic disease; cyclic.
16 Synthetic.
17
18 Key Location/Qualifiers
19 Modified-site 1..5
20 /note= "the amino group of L-Val(1) forms a peptide
21 bond with the carboxy group of L-Arg(5)"
22
23 Misc-difference 3
24 /note= "D-form residue"
25
26 Misc-difference 4
27 /note= "D-form residue"
28
29 EP596350-A.
30
31 11-MAY-1994.
32
33 23-OCT-1993. 93EP-0117204.
34
35 06-NOV-1992. 92DE-4237456.
36 (MERE ) MERCK PATENT GMBH.
37
38 Diotenbach B., Felding-Habermann B., Haubner K., Hoelzemann G.;
39 Joneczyk A., Kessler H., Melzer G., Eppmann F., Wermuth J.;
40 WPI: 1994 152792/19.
41
42 New cyclopeptide(s) of five aminoacid(s) - are useful in
43 pharmaceutical preparations and in the treatment of metabolic
44 diseases, thrombosis and heart attacks
45
46 Example 1: Page 6; 7pp; German.
47
48 This peptide is a specific example of a generic cyclic pentapeptide
49 formula. The cyclopeptides are useful for the prophylaxis and
50 treatment of metabolic diseases, thrombosis, heart attacks,
51 arteriosclerosis, inflammation, angina pectoris and tumours.
52 See AAR52716 R52720 for specifically claimed examples.
53
54 Sequence 5 AA:
55
56 Query Match 1.0%; Score 4; DB 15; Length 5;
57 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
58 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
59
60 329 VFDA 332
61 1111
62 1 VFDA 4
63
64 RESULT 33
65 AAR52731
66 ID AAR52731 standard; peptide; 5 AA.
67 AC AAR52731;
68
69 22-DEC-1994 (first entry)
70
71 Cyclic-(L-Val-L-Phe-L-Asp-D-Ala-D-Arg).
72
73 cyclopeptide; thrombosis; heart attack; arteriosclerosis;
74 inflammation; apoplexy; angina pectoris; tumour; therapy;
75 metabolic disease; cyclic.
76 Synthetic.
77
78 Key Location/Qualifiers
79 Modified-site 1..5
80 /note= "the amino group of L-Val(1) forms a peptide
81 bond with the carboxy group of D-Arg(5)"
82
83 Misc-difference 4
84 /note= "D form residue"
85
86 Misc-difference 5
87 /note= "D-form residue"
88
89 EP596350-A.

```


XX 11 MAY 1994
XX 12 MAY 1994
XX 13 MAY 1994
XX 14 MAY 1994
XX 15 MAY 1994
XX 16 MAY 1994
XX 17 MAY 1994
XX 18 MAY 1994
XX 19 MAY 1994
XX 20 MAY 1994
XX 21 MAY 1994
XX 22 MAY 1994
XX 23 MAY 1994
XX 24 MAY 1994
XX 25 MAY 1994
XX 26 MAY 1994
XX 27 MAY 1994
XX 28 MAY 1994
XX 29 MAY 1994
XX 30 MAY 1994
XX 31 MAY 1994
XX 1 JUN 1994
XX 2 JUN 1994
XX 3 JUN 1994
XX 4 JUN 1994
XX 5 JUN 1994
XX 6 JUN 1994
XX 7 JUN 1994
XX 8 JUN 1994
XX 9 JUN 1994
XX 10 JUN 1994
XX 11 JUN 1994
XX 12 JUN 1994
XX 13 JUN 1994
XX 14 JUN 1994
XX 15 JUN 1994
XX 16 JUN 1994
XX 17 JUN 1994
XX 18 JUN 1994
XX 19 JUN 1994
XX 20 JUN 1994
XX 21 JUN 1994
XX 22 JUN 1994
XX 23 JUN 1994
XX 24 JUN 1994
XX 25 JUN 1994
XX 26 JUN 1994
XX 27 JUN 1994
XX 28 JUN 1994
XX 29 JUN 1994
XX 30 JUN 1994
XX 1 JUL 1994
XX 2 JUL 1994
XX 3 JUL 1994
XX 4 JUL 1994
XX 5 JUL 1994
XX 6 JUL 1994
XX 7 JUL 1994
XX 8 JUL 1994
XX 9 JUL 1994
XX 10 JUL 1994
XX 11 JUL 1994
XX 12 JUL 1994
XX 13 JUL 1994
XX 14 JUL 1994
XX 15 JUL 1994
XX 16 JUL 1994
XX 17 JUL 1994
XX 18 JUL 1994
XX 19 JUL 1994
XX 20 JUL 1994
XX 21 JUL 1994
XX 22 JUL 1994
XX 23 JUL 1994
XX 24 JUL 1994
XX 25 JUL 1994
XX 26 JUL 1994
XX 27 JUL 1994
XX 28 JUL 1994
XX 29 JUL 1994
XX 30 JUL 1994
XX 31 JUL 1994
XX 1 AUG 1994
XX 2 AUG 1994
XX 3 AUG 1994
XX 4 AUG 1994
XX 5 AUG 1994
XX 6 AUG 1994
XX 7 AUG 1994
XX 8 AUG 1994
XX 9 AUG 1994
XX 10 AUG 1994
XX 11 AUG 1994
XX 12 AUG 1994
XX 13 AUG 1994
XX 14 AUG 1994
XX 15 AUG 1994
XX 16 AUG 1994
XX 17 AUG 1994
XX 18 AUG 1994
XX 19 AUG 1994
XX 20 AUG 1994
XX 21 AUG 1994
XX 22 AUG 1994
XX 23 AUG 1994
XX 24 AUG 1994
XX 25 AUG 1994
XX 26 AUG 1994
XX 27 AUG 1994
XX 28 AUG 1994
XX 29 AUG 1994
XX 30 AUG 1994
XX 31 AUG 1994
XX 1 SEP 1994
XX 2 SEP 1994
XX 3 SEP 1994
XX 4 SEP 1994
XX 5 SEP 1994
XX 6 SEP 1994
XX 7 SEP 1994
XX 8 SEP 1994
XX 9 SEP 1994
XX 10 SEP 1994
XX 11 SEP 1994
XX 12 SEP 1994
XX 13 SEP 1994
XX 14 SEP 1994
XX 15 SEP 1994
XX 16 SEP 1994
XX 17 SEP 1994
XX 18 SEP 1994
XX 19 SEP 1994
XX 20 SEP 1994
XX 21 SEP 1994
XX 22 SEP 1994
XX 23 SEP 1994
XX 24 SEP 1994
XX 25 SEP 1994
XX 26 SEP 1994
XX 27 SEP 1994
XX 28 SEP 1994
XX 29 SEP 1994
XX 30 SEP 1994
XX 1 OCT 1994
XX 2 OCT 1994
XX 3 OCT 1994
XX 4 OCT 1994
XX 5 OCT 1994
XX 6 OCT 1994
XX 7 OCT 1994
XX 8 OCT 1994
XX 9 OCT 1994
XX 10 OCT 1994
XX 11 OCT 1994
XX 12 OCT 1994
XX 13 OCT 1994
XX 14 OCT 1994
XX 15 OCT 1994
XX 16 OCT 1994
XX 17 OCT 1994
XX 18 OCT 1994
XX 19 OCT 1994
XX 20 OCT 1994
XX 21 OCT 1994
XX 22 OCT 1994
XX 23 OCT 1994
XX 24 OCT 1994
XX 25 OCT 1994
XX 26 OCT 1994
XX 27 OCT 1994
XX 28 OCT 1994
XX 29 OCT 1994
XX 30 OCT 1994
XX 31 OCT 1994
XX 1 NOV 1994
XX 2 NOV 1994
XX 3 NOV 1994
XX 4 NOV 1994
XX 5 NOV 1994
XX 6 NOV 1994
XX 7 NOV 1994
XX 8 NOV 1994
XX 9 NOV 1994
XX 10 NOV 1994
XX 11 NOV 1994
XX 12 NOV 1994
XX 13 NOV 1994
XX 14 NOV 1994
XX 15 NOV 1994
XX 16 NOV 1994
XX 17 NOV 1994
XX 18 NOV 1994
XX 19 NOV 1994
XX 20 NOV 1994
XX 21 NOV 1994
XX 22 NOV 1994
XX 23 NOV 1994
XX 24 NOV 1994
XX 25 NOV 1994
XX 26 NOV 1994
XX 27 NOV 1994
XX 28 NOV 1994
XX 29 NOV 1994
XX 30 NOV 1994
XX 1 DEC 1994
XX 2 DEC 1994
XX 3 DEC 1994
XX 4 DEC 1994
XX 5 DEC 1994
XX 6 DEC 1994
XX 7 DEC 1994
XX 8 DEC 1994
XX 9 DEC 1994
XX 10 DEC 1994
XX 11 DEC 1994
XX 12 DEC 1994
XX 13 DEC 1994
XX 14 DEC 1994
XX 15 DEC 1994
XX 16 DEC 1994
XX 17 DEC 1994
XX 18 DEC 1994
XX 19 DEC 1994
XX 20 DEC 1994
XX 21 DEC 1994
XX 22 DEC 1994
XX 23 DEC 1994
XX 24 DEC 1994
XX 25 DEC 1994
XX 26 DEC 1994
XX 27 DEC 1994
XX 28 DEC 1994
XX 29 DEC 1994
XX 30 DEC 1994
XX 31 DEC 1994

11 supported inactive immune and health protection of diet with
12 monotype binding proteins of receptor
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

CC AAR55746 can be used to mediate cascade kinase activity; see abstract
 CC with cell proliferation.

XX Sequence 5 AA:

Query Match 1.00; Score 4; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 SEEF 174
 IIII
 Db 2 SEEF 5

RESULT 46
 AAR54165
 ID AAR54165 standard; Protein: 5 AA.

XX AAR54165;

XX 19-FEB-1995 (first entry)

DE CHA255 heavy chain CDR3 clone 3.2.1.

XX Polymerase chain reaction primer; PCR; amplification; heavy; light;
 KW chain; complementarity determining region; CDR; variable; constant;
 KW region; monoclonal antibody; MA; binding affinity; EDTA; DETA;
 KW tumour; cancer; colorectal; breast; metal chelate; hapten.

XX Synthetic

XX AD9450602-A.

XX 26 MAY-1994.

XX 10-NOV-1993; 99AB 0059002.

XX 12-NOV-1993; 99AB 0059002.

XX (HYBR) HYBR11111111

XX Atwell PM, Moore MD;

XX WP1: 1994-209064/26.

XX N-terminus AAR54165

XX Polypeptide used in imaging and treatment of carcinomas and
 PI tumours - comprising subunit antibody CDR having binding affinity
 PI for metal chelate of EDTA or DETA or analogues

XX Claim 25; Fig 4A; 6pp; English.

XX The sequences shown in AAR54165-66 represent the wild type and
 CC mutagenised variants of the complementarity determining region 3 (CDR3)
 CC of the antibody designated CHA255. CHA255 is a murine monoclonal
 CC antibody (MAb) which is capable of binding complexes. Mutagenesis of
 CC these CDRs causes the production of polypeptides with a particularly
 CC high binding affinity for EDTA or DETA metal complexes. CDR1 and -3
 CC of the heavy chain, and CDR2 and -3 of the light chain were targeted
 CC for mutagenesis. Five residues of both CDR1 and -3 of the CHA255
 CC heavy chain (two of seven residues of light chain CDR and six of
 CC nine light chain CDR3 residues were specifically targeted for
 CC codon-based mutagenesis. The mutagenised MAB's can be used in
 CC compositions for in vivo imaging of malignant tissues or tumours. They
 CC are also useful for the treatment of malignant tissues or tumours eg.
 CC colorectal or breast cancer. Both methods involve the use of
 CC radionuclides which bind to metal chelates or haptens which are
 CC specifically delivered to the target site by a targeting molecule. CDR
 CC derived peptides may be used to construct bi-functional antibodies
 CC having dual specificities, or as donor or recipients of CDR sequences.

XX Sequence 5 AA:

Query Match 1.00; Score 4; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 RPTG 148
 IIII
 Db 2 RPTG 5

RESULT 47
 AAR54821
 ID AAR54821 standard; peptide: 5 AA.

XX AAR54821;

XX 31-OCT-1994 (first entry)

DE Lysine contg. peptide with immunomodulatory activity.

XX Immunodeficient; immunodepressed; hyperactive immunity; AIDS;

XX radiation injury; Rheumatoid arthritis; multiple sclerosis;

XX infection; atopic states; allergies; leucocyte disorders; anaemia;

XX Synthetic

XX W09409804-A.

XX 11-MAY-1994.

XX 28-OCT-1993; 93WO-0510441.

XX 29-OCT-1993; 93US-0967633.

XX (CYTO) CYTOAVER INT NV.

XX Green LR, Young VT, Phallingson VP, Mikhaltsov AN;

XX Mikhaltsov AN, Mikhaltsov AN, Mikhaltsov AN, Mikhaltsov AN;

XX WP1: 1994-15/115/20.

XX Lysine contg. peptide(s), including new cpds., with
 PI immunomodulatory activity - for treating defective or hyperactive
 PI immune states, treatment or prevention of infection, atopic
 PI states and leukocytic disorders, also vaccine adjuvants

XX Claim 4; Page 49; 60pp; English.

XX The sequence is a claimed example of the generic peptide R1-Glx-Glx-
 CC Lys R2 (where Glx = Glu/Gln and R1 and R2 are peptide sequences of
 CC less than 7 residues or hydrogen). The generic peptide has 5-9 amino
 CC acid residues in total. Such peptides are used to treat various
 CC immunorestates. Specifically they are used in human or veterinary
 CC medicine to modulate the activity of the host's immune system, for
 CC treating infections, for treating atopic states of leukocytic
 CC disorders and as adjuvants for vaccination. They can also be used
 CC to treat anaemias.

XX See also AAR54820-27.

XX Sequence 5 AA:

Query Match 1.00; Score 4; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 AEPK 272
 IIII
 Db 2 AEPK 5

RESULT 48
 AAR77331
 ID AAR77331 standard; Peptide: 5 AA.

FT pyridyl)piperazine, 1- γ -glutamate-4-(2-methoxyphenyl)piperazine, 1- γ -glutamate-4-(4-methoxyphenyl)piperazine, opt. D-form residue"
 FT Misc-difference 4 /note- "D-form residue"
 FT Modified-site 5 /note- "joined via a peptide linkage to Arg1"
 XX
 CS Synthetic.
 XX W99634012 A1.
 XX 31-OCT-1996.
 XX 25-APR-1996; 96WO-JP01140.
 XX 09-MAY-1995; 95JP-0110933.
 XX 28-APR-1995; 95JP-0106775.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Furuya S, Kato K, Kitada G.
 XX WPI: 1996-497569/49
 XX Cyclic pentapeptide(s), some new, as LH-RH receptor antagonists -
 PT used to treat or prevent sex hormone-dependent disorders, e.g.
 PT cancer, also for control of pregnancy and menstruation and to
 PT improve meat quality in animals
 XX Example 9 17, Page 71, 1996, English.
 XX This peptide represents a cyclic peptide which is included in the
 CC intensifying hormone releasing hormone (LH RH) receptor antagonist
 CC composition of the invention. Peptides such as this are used to
 CC prevent or treat sex hormone dependent disorders in human or veterinary
 CC medicine, esp. cancers (of the prostate, uterus, breast or pituitary),
 CC prostatic hyperplasia, endometriosis, hyperandrogenic puberty, hot
 CC also amenorrhea, premenstrual syndrome, multifactorial ovarian syndrome,
 CC comedo, etc. also to control pregnancy (radiation for men or women,
 CC also to induce ovulation) and the menstrual cycle. They are also used
 CC to control oestrus in animals, to improve meat quality and control
 CC growth, and to promote spawning in fish. They may also inhibit the
 CC transient increase in testosterone blood levels caused by admin of
 CC superagonists such as leuprolin acetate
 XX Sequence 5 AA;
 SQ
 Query Match 1.0%; Score 4; DB 17; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 EYWL 359
 DB 2 EYWL 5
 RESULT 41
 AAR98672
 ID AAR98672 standard; peptide; 5 AA.
 AC AAR98672;
 XX 11-MAR-1997 (first entry)
 DE Peptide 52 from 88 member diverse mimotope panel.
 XX Panel, mimotopes, decreasing hydrophobicity, periodic variation;
 FW hydrophobic moment; antibody repertoire, identification, 100;
 KW candidate, receptor, binding, liquid, rational, design, selection;
 FW treatment; tumour; production; immunological reagent; analyte;
 KW detection; trace contaminant; mimotope.

XX CS Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note "acylated"
 XX US5541070-A.
 XX 30-JUL-1996.
 XX 13-OCT-1987; 87US-0108130.
 XX 06-DEC-1989; 89US-0447009.
 XX 13-OCT-1987; 87US-0108130.
 XX 11-OCT-1988; 88US-0255906.
 XX 08-SEP-1993; 93US-0118133.
 XX (KAUV/) KAUVAR L M.
 XX Kaurar LM;
 XX WPI: 1996-461957/36.
 XX Identifying candidate drugs that bind a specific receptor by
 PT competitive reaction with panel of mimotopes(s) - useful in rational
 PT drug design
 XX Example 3, Fig 3, 27pp, English.
 XX The present peptide is a member of a panel of 88 pentapeptide
 CC mimotopes designed on the basis of decreasing hydrophobicity, and
 CC periodic variation of hydrophobic moment. The panel was synthesised
 CC using the method of Gysin, H. M., et al, Proc. Natl Acad Sci
 CC USA (1984), which uses sets of 96 pins, the remaining 8
 CC polyethylene pins being controls. The mimotopes were then mixed,
 CC 2:5:1 labelled and tested with individual members of a basal
 CC antibody (Ab) repertoire. Nearly uniform binding to all members was
 CC found. The test was then repeated with the addition of a defined
 CC amt. of analyte. A small number showed greatly increased labelling,
 CC these Ab representing the successful result of an initial screen
 CC for those that pref. bind analyte.
 CC The above is an example of a claimed method for identifying members
 CC of a panel of candidate drugs, that bind to a receptor having a
 CC known ligand. It is useful in rational drug design, e.g. selection
 CC of monoclonal Ab for treating individual tumours, and for the prodn.
 CC of immunological reagents for any analyte, including those not
 CC normally detectable by immunoassay, e.g. trace contaminants in soil,
 CC air or water.
 XX Sequence 5 AA;
 SQ
 Query Match 1.0%; Score 4; DB 17; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 362 NSWG 365
 DB 1 NSWG 4
 RESULT 42
 AAW03248
 ID AAW03248 standard; peptide; 5 AA.
 XX AAW03248;
 AC AAW03248;
 XX 16-MAR-1997 (first entry)
 DE Peptide isolated from soybean hydrolysate.
 XX Peptide, ampicillin inhibiting enzyme inhibitor; ACE.
 XX


```

QV 161 KSSR 164
DE 1111
DB 2 KSSR 5

RESULT 49
AAW51412
ID AAW51412 standard, Peptide, 5 AA.
XX
AC AAW51412:
DE 12-OCT-1998 (first entry)
DE Human protease activated receptor 3 (PAR3) peptide (aa48-52).
XX
KW Protease-activated receptor 3, PAR3, thrombin receptor; human;
KW 5-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; inflammation;
XX
OS Homo sapiens.
XX
FN W9618456-A1.
XX
DE 07-MAY 1998.
XX
DE 29-OCT-1997: 97W-0510742.
XX
DE 30-OCT 1996: 96OS-0742440.
XX
DE (RECD ) UNIV CALIF-RNFA.
XX
PI Connolly A. Orophilin SR. Ishihara B;
XX
DE WPI: 1998-271905/24
XX
DE DNA encoding protease-activated receptor 3 - for detection of
PI specific agonists and antagonists, potentially useful for treating
PI e.g. thrombosis, atherosclerosis, inflammation etc.
XX
PS Example 1: Page 24, 74pp; English.
XX
CC This peptide corresponds to amino acid residues 48-52 of human
CC protease-activated receptor 3 (PAR3) (see AAW51406), a cell surface
CC protein which is specifically activated by thrombin. This motif
CC shows homology to the C-terminal tail (see AAW51409) of hirudin
CC and it is suggested that it docks with thrombin's fibrinogen
CC binding exosite while amino acids 45-48 (see AAW51413) dock with
CC thrombin's active centre leading to cleavage of the K38-T39 peptide
CC bond. Peptides homologous to this tethered domain of PAR3 (see
CC AAW51414-15) may be potential agonists of PAR3 activity. PAR3 is
CC a novel thrombin receptor and can be used to screen for agonists
CC and antagonists of thrombin useful e.g. for treatment of
CC thrombosis, atherosclerosis and inflammation.
XX
SQ Sequence 5 AA.

Query Match 1.0%; Score 4; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 86 FEEF 89
DE 1111
DB 1 FEEF 4

RESULT 50
AAW56410
ID AAW56410 standard; peptide; 5 AA.
XX
AC AAW56410:
XX
DE 05-AUG-1998 (first entry)
DE Nuclear localisation signal (NLS).
XX
KW Signal peptide; nuclear localisation signal; NLS;
KW immunosuppressive activity; inhibitor; nuclear translocation inhibitor;
KW nuclear translocation, treatment, immune disorder, autoimmune disease;
KW hypersensitivity; sepsis, prevention, septic shock, antiviral agent;
KW tumour growth suppressor.
XX
OS unidentified.
XX
FN W9811907-A.
XX
DE 26 MAR-1998.
XX
DE 15-SEP-1997: 97W-0516217.
XX
DE 12-JUL 1997: 97J3-0528958.
XX
DE 20-JUL-1996: 96J3-0026978.
XX
DE (HRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Blake J. Cleaveland JS, Hattar CK, Nadler SG;
XX
DE WPI: 1998-217028/19.
XX
PI Nuclear translocation inhibitor polypeptides - comprising signal
PI sequence for delivery through the cytoplasmic membrane and at least
PI 2 nuclear localisation sequences
XX
DE claim 10; Page 44; 69pp; English.
XX
CC Peptides AAW56400-13 represent nuclear localisation signals (NLS). The
CC sequence is used to construct the nuclear translocation inhibitor
CC polypeptide of the invention. Nuclear translocation inhibitor
CC polypeptides comprise a signal sequence peptide capable of delivering
CC the polypeptide through the cytoplasmic membrane into a cell, and at
CC least 2 NLSs. The polypeptides can be used to inhibit nuclear
CC translocation of a cellular protein. In addition, since the nuclear
CC translocation of certain cellular peptides is required for the host
CC organism to mount an immune response, the polypeptide inhibitors are
CC useful as immunosuppression agents. The polypeptides can therefore be
CC used for the treatment of immune disorders including autoimmune
CC diseases. The polypeptides can also be used for treating physical
CC symptoms manifested by responses to allergens which can initiate a state
CC of hypersensitivity, for the treatment of rashes and in the prevention
CC of septic shock, antiviral agents, tumour growth suppressors, and for
CC transcriptionally modulating the expression of cellular genes.
XX
SQ Sequence 5 AA;

Query Match 1.0%; Score 4; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 90 KKKY 93
DE 1111
DB 1 KKKY 4

Search completed: October 9, 2002, 11:53:08
Job time : 40 secs

```


GenCore version 5.1.3
Copyright (c) 1996-2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:52:21, Search time 22 seconds
(without alignments)
1751.446 Million cell updates/sec

Title: 02-09-09-062-4
Perfect score: 401
Sequence: 1 MDIGKRVHHQHWISCHYIA.....KKKCGHVLKPVYVINGSI 101

Scoring table: OLIGO
Gapop 60 0 0 Gapext 60 0

Searched: 29138 seqs, 9609324 residues

Word size: 0

Total number of hits satisfying chosen parameters: 515

Minimum DB seq length: 0
Maximum DB seq length: 6

Post processing: Listing first 1000 summaries

Database: PIR71:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	1-2	6	ps0009	angiotensin-conver
2	1	1-0	6	170516	T-cell receptor be
3	3	0-2	3	130216	angiotensin-conver
4	3	0-2	4	318401	glycyl-leucyl-gly
5	3	0-2	4	170505	hypothetical prote
6	3	0-2	4	161899	protamine PI - oia
7	3	0-2	4	170655	T-cell receptor be
8	3	0-2	4	137012	protamine PI - Cor
9	3	0-2	4	194199	protamine PI - sar
10	3	0-2	5	160521	glycogen phosphory
11	3	0-2	5	376151	99F2 protein - Xan
12	3	0-2	5	271836	99F2 protein - RISK
13	3	0-2	5	135561	ribosomal protein
14	3	0-2	5	139956	ribosomal protein
15	3	0-2	5	139969	ribosomal protein
16	3	0-2	5	960271	major protein, andi
17	3	0-2	5	122552	5-phosphorylchit-9a
18	3	0-2	5	122552	angiotensin-conver
19	3	0-2	5	122552	photosystem - 10.4
20	3	0-2	5	951077	alpha amylase - ri
21	3	0-2	5	144545	tubulin - giant: Af
22	3	0-2	5	150385	myosin light chain
23	3	0-2	5	370520	lg kappa chain V-1
24	3	0-2	5	352983	serominal plasma pro
25	3	0-2	5	170610	T-cell receptor be
26	3	0-2	5	170660	T-cell receptor be
27	3	0-2	5	170661	T-cell receptor be
28	3	0-2	5	170665	T-cell receptor be
29	3	0-2	5	170577	T-cell receptor be

T-cell receptor be
T-cell receptor be
alpha-1,4-dehydrogen
peptidyl-dipeptida
halo-toxin - Pseud
hydrogensulfite re
jacalin beta-II ch
anticeplastic gly
locustakinin - mig
tubulin beta-3 cha
protamine PI - cor
alpha-1,4-glycan-p
alpha-tubulin - Ch
T cell receptor be
T-cell receptor be
T-cell receptor be
laminin B1 - weste
neuropeptide GNPR
pro-kinin 2 - pea
growth-modulating
barstin - chicken
spinal cord peptid
spinal cord peptid
thyrotropin-releas
k-phycocerythrin al
TRH like tripeptid
bradykinin-potentl
bradykinin-potentl
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
blood cell protein
cytochrome-c oxida
antho Rfamide neur
phagocytosis-stimu
tyrosine-melanosyl
cardio-excitatory n
carbon monoxide de
phenol 3-monooxyge
gamma subunit of p
biotin A - Citrob
22G-subcellular D
transmembrane hydrol
protein D - Escher
hypothetical prote
hypothetical prote
cell surface adhes
phospholipase C (E
endoalucanase P -
hypothetical prote
glucanase IV alpha
saturation-induced
ribos-4-1: protein
glucan 1,4-alpha-g
hypothetical prote
neuropeptide Antho
neuropeptide Antho
antho-PFamide neu
PMAamide - polych
PMAamide-related neu
actinin-1 - giant
metallothionein-A
myosin-light-chain
lg heavy chain: CRD
lg heavy chain: CRD
lg m chain V reql
lg m chain V reql
synaptosomal-assoc
T-cell receptor be
T-cell receptor be

249	2	0.5	6	2	376764	ribosomal protein
250	2	0.5	6	2	A19789	transferrin - bovi
251	2	0.5	6	2	P43302	chey glycoprotein
252	2	0.5	6	2	A19792	acylaminopropyl-pip
253	2	0.5	6	2	A11856	fatty acid synthase
254	2	0.5	6	2	B45640	vertebral degener
255	2	0.5	6	2	A49900	P 372.5 residues
256	2	0.5	6	2	A46474	pc epsilon klib - m
257	2	0.5	6	2	B32923	in mu chain P-regi
258	2	0.5	6	2	B7345	MHC H2-K k cell su
259	2	0.5	6	2	B65546	MHC H2-L antigen -
260	2	0.5	6	2	B59142	platelet-derived q
261	2	0.5	6	2	B70629	T cell receptor be
262	2	0.5	6	2	B70511	T cell receptor be
263	2	0.5	6	2	B70511	T cell receptor be
264	2	0.5	6	2	B70511	T cell receptor be
265	2	0.5	6	2	B70511	T cell receptor be
266	2	0.5	6	2	B70511	T cell receptor be
267	2	0.5	6	2	B70519	T cell receptor be
268	2	0.5	6	2	B70604	T cell receptor be
269	2	0.5	6	2	B70640	T cell receptor be
270	2	0.5	6	2	B70643	T cell receptor be
271	2	0.5	6	2	B70637	T cell receptor be
272	2	0.5	6	2	B70621	T cell receptor be
273	2	0.5	6	2	B70616	T cell receptor be
274	2	0.5	6	2	B70605	T cell receptor be
275	2	0.5	6	2	B70618	T cell receptor be
276	2	0.5	6	2	B70697	T cell receptor be
277	2	0.5	6	2	B70619	T cell receptor be
278	2	0.5	6	2	B70511	T cell receptor be
279	2	0.5	6	2	B70652	T cell receptor be
280	2	0.5	6	2	B70657	T cell receptor be
281	2	0.5	6	2	B70652	T cell receptor be
282	2	0.5	6	2	B70668	T cell receptor be
283	2	0.5	6	2	B70658	T cell receptor be
284	2	0.5	6	2	B70643	T cell receptor be
285	2	0.5	6	2	B70543	T cell receptor be
286	2	0.5	6	2	B70560	T cell receptor be
287	2	0.5	6	2	B70559	T cell receptor be
288	2	0.5	6	2	B70723	T cell receptor be
289	2	0.5	6	2	B70718	T cell receptor be
290	2	0.5	6	2	B70589	T cell receptor be
291	2	0.5	6	2	B70587	T cell receptor be
292	2	0.5	6	2	B70568	T cell receptor be
293	2	0.5	6	2	B70673	T cell receptor be
294	2	0.5	6	2	B70593	T cell receptor be
295	2	0.5	6	2	B70726	T cell receptor be
296	2	0.5	6	2	B70727	T cell receptor be
297	2	0.5	6	2	B70703	T cell receptor be
298	2	0.5	6	2	B70715	T cell receptor be
299	2	0.5	6	2	B70738	T cell receptor be
300	2	0.5	6	2	A41946	T cell receptor ga
301	2	0.5	6	2	A41946	T cell receptor ga
302	2	0.5	6	2	B54524	cytochrome T lympho
303	2	0.5	6	2	B71349	beta crystallin B2
304	2	0.5	6	2	A35039	hypothalamic cellu
305	2	0.5	6	2	B70564	hypothetical prot
306	2	0.5	6	2	B70556	orf 3 para 5'-regi
307	1	0.2	3	3	B70596	thryoliberin - por
308	1	0.2	3	3	B70596	thryoliberin - por
309	1	0.2	3	3	B70596	thryoliberin - por
310	1	0.2	3	3	B70596	thryoliberin - por
311	1	0.2	3	3	B70596	thryoliberin - por
312	1	0.2	3	3	B70596	thryoliberin - por
313	1	0.2	3	3	B70596	thryoliberin - por
314	1	0.2	3	3	B70596	thryoliberin - por
315	1	0.2	3	3	B70596	thryoliberin - por

ALIGNMENTS

P00008	angiotensin converting enzyme inhibitor (ELP 1)	common lig
N.Alternate names:	ficus latex peptide 1	
C.Species:	Ficus carica (common fig)	
C.Date:	07 Jun 1990 #sequence_revision 07 Jun 1990 #text_change 08 Dec 1995	
C.Accession:	P00008	
R.Mutagenesis:	S. Miyoshi, S. Tanaka, H. Arita, Biol. Chem. 263, 2763-2767, 1989	
A.Title:	Angiotensin 1-converting enzyme inhibitors derived from Ficus carica.	
A.Reference:	P00008	
A.Accession:	P00008	
A.Molecule type:	protein	
A.Residues:	1-6 <MAR>	
A.Experimental source:	latex	
C.Keywords:	angiotensin converting enzyme inhibitor	
Query Match:	1.00, Score 5, DB 2, Length 6;	
Best Local Similarity:	100.00, Pred. No. 2.8e+05,	
Matches:	5, Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	188 VNP1R 192	
DB	2 VNP1R 6	
RESULT 2		
T cell receptor beta chain V-E-J region (100-4AP) - mouse (fragment)		
C.Species:	Mus musculus (house mouse)	
C.Date:	17 Jul 1992 #sequence_revision 17 Jul 1992 #text_change 20 May 1997	
C.Accession:	PT0516	
R.February, A.J.		
C.Date:	Mod 174, 115-124, 1991	
A.Title:	Functional sequences of fetal T cell receptor beta chains have few N regions	
A.Reference:	number: PT0516, MIM: 41277601	
A.Accession:	PT0516	
A>Status:	translation not shown	
A.Molecule type:	RNA	
A.Residues:	1-6 <PFE>	
A.Experimental source:	adult thymus, strain BALB/c	
C.Keywords:	T cell receptor	
Query Match:	1.00, Score 4, DB 2, Length 6;	
Best Local Similarity:	100.00, Pred. No. 2.8e+05,	
Matches:	4, Conservative 0, Mismatches 0, Indels 0, Gaps 0;	
QY	369 GKG 372	
DB	3 GKG 6	
RESULT 3		
P00010	angiotensin converting enzyme inhibitor (ELP 3)	common lig
N.Alternate names:	ficus latex peptide 3	
C.Species:	Ficus carica (common fig)	
C.Date:	15 Jun 2001 #sequence_revision 15 Jun 2001 #text_change 15 Jun 2001	
C.Accession:	P00010	
R.Mutagenesis:	S. Miyoshi, S. Tanaka, H. Arita, Biol. Chem. 263, 2763-2767, 1989	
A.Title:	Angiotensin 1-converting enzyme inhibitors derived from Ficus carica.	
A.Reference:	P00010	
A.Accession:	P00010	
A.Molecule type:	protein	
A.Residues:	1-3 <MAR>	
A.Experimental source:	latex	
C.Keywords:	angiotensin converting enzyme inhibitor	
Query Match:	0.78, Score 3, DB 3, Length 3;	
Best Local Similarity:	100.00, Pred. No. 2.8e+05;	
Matches:	3, Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	359 LVR 361	

13. [REDACTED]
14. [REDACTED]
15. [REDACTED]
16. [REDACTED]
17. [REDACTED]
18. [REDACTED]
19. [REDACTED]
20. [REDACTED]
21. [REDACTED]
22. [REDACTED]
23. [REDACTED]
24. [REDACTED]
25. [REDACTED]
26. [REDACTED]
27. [REDACTED]
28. [REDACTED]
29. [REDACTED]
30. [REDACTED]
31. [REDACTED]
32. [REDACTED]
33. [REDACTED]
34. [REDACTED]
35. [REDACTED]
36. [REDACTED]
37. [REDACTED]
38. [REDACTED]
39. [REDACTED]
40. [REDACTED]
41. [REDACTED]
42. [REDACTED]
43. [REDACTED]
44. [REDACTED]
45. [REDACTED]
46. [REDACTED]
47. [REDACTED]
48. [REDACTED]
49. [REDACTED]
50. [REDACTED]
51. [REDACTED]
52. [REDACTED]
53. [REDACTED]
54. [REDACTED]
55. [REDACTED]
56. [REDACTED]
57. [REDACTED]
58. [REDACTED]
59. [REDACTED]
60. [REDACTED]
61. [REDACTED]
62. [REDACTED]
63. [REDACTED]
64. [REDACTED]
65. [REDACTED]
66. [REDACTED]
67. [REDACTED]
68. [REDACTED]
69. [REDACTED]
70. [REDACTED]
71. [REDACTED]
72. [REDACTED]
73. [REDACTED]
74. [REDACTED]
75. [REDACTED]
76. [REDACTED]
77. [REDACTED]
78. [REDACTED]
79. [REDACTED]
80. [REDACTED]
81. [REDACTED]
82. [REDACTED]
83. [REDACTED]
84. [REDACTED]
85. [REDACTED]
86. [REDACTED]
87. [REDACTED]
88. [REDACTED]
89. [REDACTED]
90. [REDACTED]
91. [REDACTED]
92. [REDACTED]
93. [REDACTED]
94. [REDACTED]
95. [REDACTED]
96. [REDACTED]
97. [REDACTED]
98. [REDACTED]
99. [REDACTED]
100. [REDACTED]

101. [REDACTED]
102. [REDACTED]
103. [REDACTED]
104. [REDACTED]
105. [REDACTED]
106. [REDACTED]
107. [REDACTED]
108. [REDACTED]
109. [REDACTED]
110. [REDACTED]
111. [REDACTED]
112. [REDACTED]
113. [REDACTED]
114. [REDACTED]
115. [REDACTED]
116. [REDACTED]
117. [REDACTED]
118. [REDACTED]
119. [REDACTED]
120. [REDACTED]
121. [REDACTED]
122. [REDACTED]
123. [REDACTED]
124. [REDACTED]
125. [REDACTED]
126. [REDACTED]
127. [REDACTED]
128. [REDACTED]
129. [REDACTED]
130. [REDACTED]
131. [REDACTED]
132. [REDACTED]
133. [REDACTED]
134. [REDACTED]
135. [REDACTED]
136. [REDACTED]
137. [REDACTED]
138. [REDACTED]
139. [REDACTED]
140. [REDACTED]
141. [REDACTED]
142. [REDACTED]
143. [REDACTED]
144. [REDACTED]
145. [REDACTED]
146. [REDACTED]
147. [REDACTED]
148. [REDACTED]
149. [REDACTED]
150. [REDACTED]
151. [REDACTED]
152. [REDACTED]
153. [REDACTED]
154. [REDACTED]
155. [REDACTED]
156. [REDACTED]
157. [REDACTED]
158. [REDACTED]
159. [REDACTED]
160. [REDACTED]
161. [REDACTED]
162. [REDACTED]
163. [REDACTED]
164. [REDACTED]
165. [REDACTED]
166. [REDACTED]
167. [REDACTED]
168. [REDACTED]
169. [REDACTED]
170. [REDACTED]
171. [REDACTED]
172. [REDACTED]
173. [REDACTED]
174. [REDACTED]
175. [REDACTED]
176. [REDACTED]
177. [REDACTED]
178. [REDACTED]
179. [REDACTED]
180. [REDACTED]
181. [REDACTED]
182. [REDACTED]
183. [REDACTED]
184. [REDACTED]
185. [REDACTED]
186. [REDACTED]
187. [REDACTED]
188. [REDACTED]
189. [REDACTED]
190. [REDACTED]
191. [REDACTED]
192. [REDACTED]
193. [REDACTED]
194. [REDACTED]
195. [REDACTED]
196. [REDACTED]
197. [REDACTED]
198. [REDACTED]
199. [REDACTED]
200. [REDACTED]

A:Cross references: EMBL:Z12147; NID:q10114; FIDN:AA70131.1; FID:q179349

Query Match 0.78; Score 3; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MAR 145

DB 1 MAR 4

RESULT 10

A60521

glycogen phosphorylase (EC 2.4.1.1), muscle mullet (*Liza ramada*) (fragment)

NCBI-entrez name: glycogen phosphorylase b

C:Species: *Liza ramada*

C:Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text_change 11-May-2000

C:Accession: A60521

R:Recombinant: I. J. Radante, I.V.

C:Comp. Biochem. Physiol. B 95, 295-301, 1990

A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus

A:Reference number: A60521; MUID:90227907

A:Accession: A60521

A:Molecule type: protein

A:Residues: 1-5 <RES>

C:Superfamily: phosphorylase

C:Keywords: glycolysis;transcription; beta-actin;transcription; phosphorylase

F:Binding site: phosphate (2x) (nonident) (17 phosphorylase b kinase) #status expert

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 ISV 415

DB 2 ISV 4

RESULT 11

S70154

URE2 protein - *Xanthomonas* sp.

C:Species: *Xanthomonas* sp.

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999

C:Accession: S70154

R:Recombinant: G.Y.; Minamide, F.Z.; Bass, F.A.; Variana, G.V.; Mithiboin, S.V.; Kiklis, P.

Mol. Microbiol. 17, 1189-1200, 1995

A:Title: Four genes, two ends, and a res region are involved in transposition of *Tn*5034

A:Reference number: S70154; MUID:96130850

A:Accession: S70154

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: RNA

A:Residues: 1-5 <RES>

A:Cross-references: EMBL:J10672; FIDN:AA7029.1; FID:q14599

A:Notes: the nucleotide sequence was submitted to the EMBL data library, May 1995

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 SAS 168

DB 2 SAS 4

RESULT 12

B41846

20K protein - *Rickettsia rickettsii* (fragment)

C:Species: *Rickettsia rickettsii*

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 28-May-1999

C:Accession: B41846

R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.

J. Bacteriol. 170, 4493-4500, 1988

A:Title: Expression of the gene encoding the 17-kilodalton antigen from *Rickettsia rick*

A:Reference number: A91895; MUID:W908059

A:Accession: B31846

A:Molecule type: RNA

A:Residues: 1-5 <AMD>

A:Cross references: GIM:J03371; NID:q152455; FIDN:AA015030.1; FID:q4423874

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 DTN 354

DB 2 DTN 4

RESULT 14

I39964

ribosomal protein S4 - *Bacillus circulans* (fragment)

C:Species: *Bacillus circulans*

C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

C:Accession: I39964

R:Grundy, F.J.; Henkin, T.M.

J. Bacteriol. 174, 6763-6770, 1992

A:Title: Characterization of the *Bacillus subtilis* rpsD regulatory target site.

A:Reference number: I39963; MUID:94015735

A:Accession: I39964

A>Status: preliminary; translated from GCEMM4/JDDBJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross references: GIM:M99041; NID:q143471

C:Genetics:

A:Gene: rpsD

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MAR 145

DB 1 MAR 3

RESULT 14

I39966

ribosomal protein S4 - *Bacillus licheniformis* (fragment)

C:Species: *Bacillus licheniformis*

C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

C:Accession: I39966

R:Grundy, F.J.; Henkin, T.M.

J. Bacteriol. 174, 6763-6770, 1992

A:Title: Characterization of the *Bacillus subtilis* rpsD regulatory target site.

A:Reference number: I39963; MUID:94015735

A:Accession: I39966

A>Status: preliminary; translated from 35/FMD/JDDBJ

A:Molecule type: RNA

A:Residues: 1-5 <RES>

A:Cross references: GIM:M99043; NID:q143475

C:Genetics:

A:Gene: rpsD

Query Match

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MAR 145

DB 1 MAR 3

RESULT 15

I39965

ribosomal protein S4 - *Bacillus megaterium* (fragment)

C:Species: *Bacillus megaterium*

Tullius, - giant African snail
 C:Species: Archamia fulica (giant African snail)
 C:Date: 23-Mar-1995 #sequence_revision 05 Apr 1995 #text_change 11 Jul-1997
 C:Accession: A44692
 Kobata, R., Park, C. L., Takai, T., Shiohatake, Y., Yasuda Kumakura, Y., Minakata, H.: No
 Biochem Biophys Res Commun 178, 486-493, 1991
 A:Title: Full-length cDNA of a novel neuropilin-like protein containing a D-amino acid residue isolated from
 A:Reference number: A44692; MIM:6115471
 A:Accession: A44692
 A:Molecule type: protein
 A:Residues: 1-5 <SH>
 C:Keywords: amidated carboxylated D-amino acid; neuropilin
 F:27 Modified site: 2 aspartate (Asp) #status experimental
 F:5 Modified site: 5 aspartate (Asp) and (Val) #status experimental

 Query Match 0.7% Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 142 NEF 144
 DB 11
 LE 2 NEF 4

 RESULT 22
 150485
 m708in light chain 2 - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C:Date: 14-Sep-1996 #sequence_revision 13 Sep 1996 #text_change 28 Feb 1997
 C:Accession: J10485
 R:Shen, R.; Goswami, S.K.; Mascareno, E.; Kumar, A.; Siddiqui, M.A.Q.
 Mol Cell Biol 11, 1676-1685, 1991
 A:Title: Tissue specific transfection of the human myosin light chain 2 gene is regu
 A:Reference number: J10485; MIM:61141519
 A:Accession: J10485
 A:Status: preliminary; translated from GE/EMBL/4893
 A:Molecule type: DNA
 A:Residues: 1-5 <SH>
 A:Protein reference: GE/M6396; MIM:421232
 C:Gene: MYH2

 Query Match 0.7% Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 VSA 147
 DB 11
 LE 2 VSA 4

 RESULT 23
 J10520
 14 kappa chain V-111 region (SH) - human (fragment)
 C:Species: Homo sapiens (human)
 C:Date: 24-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 16-Aug-1996
 C:Accession: J10520
 R:Anker, R.; Conley, M.E.; Pollok, B.A.
 J. Exp. Med. 169, 2109-2119, 1989
 A:Title: Clonal diversity in the B-cell repertoire of patients with X linked agammaglobu
 A:Reference number: J10511; MIM:30279157
 A:Accession: J10520
 A:Molecule type: mRNA
 A:Residues: 1-5 <SH>
 A:Note: the sequence shown here is one of eight productive V D-J mu chain rearrangements
 A:Note: a stop codon terminates the sequence in the V region
 C:Keywords: heterodimer; immunoglobulin
 F:1-5/chain: V kappa region <VRE>

 Query Match 0.7% Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 142 NEF 144
 DB 11
 LE 2 NEF 4

QY 15 SGP 17
 DB 3 SGP 5

 RESULT 24
 S62883
 seminal plasma protein II - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 28 Oct 1996 #sequence_revision 13-Mar 1997 #text_change 17-Mar-1999
 C:Accession: S62883
 R:Pearce, A.; Vardell, F.F.; Sauer, J.; Toepfer Petersen, E.; Calverley, Y.J.
 Exp Cell Res 262, 15-17, 1996
 A:Title: Crystallization and preliminary X-ray diffraction analysis of boar seminal p
 A:Reference number: S62882; MIM:96196555
 A:Accession: S62883
 A:Molecule type: protein
 A:Residues: 1-5 <SH>
 C:Complex: heterodimer; seminal plasma protein I and seminal plasma protein II
 C:Keywords: glycoprotein; heterodimer; semen

 Query Match 0.7% Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 ING 24
 DB 3 ING 5

 RESULT 25
 PT0610
 T cell receptor beta chain V D-J region (136-23) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul 1995 #sequence_revision 17 Jul 1995 #text_change 20 May-1997
 C:Accession: PT0610
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of total T cell receptor beta chains have few N regions
 A:Reference number: PT0509; MIM:91277601
 A:Accession: PT0610
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <PE>
 A:Note: internal source; newborn thymus, strain B6.D-2
 C:Keywords: T cell receptor

 Query Match 0.7% Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 168 SES 170
 DB 3 SES 5

 RESULT 26
 PT0660
 T cell receptor beta chain V-D-J region (121-181) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17 Jul 1992 #sequence_revision 17 Jul 1992 #text_change 30 May 1997
 C:Accession: PT0660
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of total T cell receptor beta chains have few N regions
 A:Reference number: PT0509; MIM:91277601
 A:Accession: PT0660
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <PE>
 A:Note: internal source; day 4 postnatal thymus, strain BALB/c
 C:Keywords: T cell receptor

alcohol dehydrogenase (NADH-dependent) class III low affinity form - cod (Gadus sp.) (fragment)
 C:Species: Gadus sp. (cod)
 C:Date: 19-Mar-1997 #sequence_revision 29-Apr-1997 #text_change 12-Jun-1998
 C:Accession: S66195
 F:Helmsvist, L.; Backström, M.; Shafiqat, J.; Danielsson, O.; Lida, J.; Hendrickson, R.C.;
 FEBS Lett. 367: 247-249, 1995
 A:Title: Multiplicity of N terminal structures of medium chain alcohol dehydrogenases, M
 enzyme.
 A:Reference number: S66195, M01D:95331482
 A:Accession: S66195
 A:Molecule type: protein
 A:Residues: 1-6 <HIS>
 C:Superfamily: alcohol dehydrogenase, long chain alcohol dehydrogenase homology
 C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 0.7% Score 3; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 2-5 AVA 207

DB 2 AVA 4

RESULT 35
 JN8651
 peptidyl-dipeptidase A inhibitory peptide (C11 - striped bonito)
 C:Species: Sarda orientalis (striped bonito)
 C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
 C:Accession: JN8651
 F:Matsumura, N.; Fujii, M.; Iakeda, Y.; Shimizu, T.
 Ploest. Biotechnol. Biochem. 57: 1743-1744, 1993
 A:Title: Isolation and characterization of antidiarrhoeal secretory enzyme inhibitor pe
 A:Reference number: JN8651, M01D:94280336
 A:Accession: JN8651
 A:Molecule type: protein
 A:Residues: 1-6 <HIS>
 C:Superfamily: bradykinin-potentiating peptide
 C:Keywords: antidiarrhoeal secretory enzyme inhibitor

Query Match 0.7% Score 3; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 3-4 VYP 395

DB 2 VYP 4

RESULT 34
 A61049
 halo-toxin - pseudomonas syringae pv. mori
 C:Species: Pseudomonas syringae pv. mori
 A:Note: host mulberry tree
 C:Date: 10-Mar-1994 #sequence_revision 18-Mar-1994 #text_change 21-Jan-1997
 C:Accession: A61049
 F:Kajimato, T.; Yokoyama, K.; Yabiro, K.; Umeda, T.; Shoji, S.; Kubota, Y.; Shibata, M.;
 Chem. Lett. 00: 679-680, 1989
 A:Title: Structure of halo-toxin, produced by phytopathogenic bacterium, Pseudomonas syri
 A:Reference number: A61049
 A:Accession: A61049
 A:Molecule type: protein
 A:Residues: 1-6 <KAT>
 C:Note: sequence confirmed by synthesis
 C:Comment: This toxin is one of the studied plant of halo bright disease in mulberry
 C:Keywords: toxin

Query Match 0.7% Score 3; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 309 GPI 311
 DB 4 GPI 6
 RESULT 35
 S11556
 hydrogensulfite reductase (EC 1.8.99.3) chain 2 - desulfovibrio thermophilus (fragment)
 N:Alternative names: bisulfite reductase, desulfotasecin
 C:Species: Desulfovibrio thermophilus
 C:Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
 C:Accession: S11556
 F:Paquet, C.; Litz, A.F.; Czechowski, M.; Kani, L.; DerVartanian, E.V.; Moura, J.J.G.;
 Biochim. Biophys. Acta 1049: 112-118, 1990
 A:Title: Purification and characterization of bisulfite reductase (desulfotasecin) f
 A:Reference number: S11024, M01D:9035276
 A:Accession: S11556
 A:Molecule type: protein
 A:Residues: 1-6 <PAD>
 C:Keywords: oxidoreductase

Query Match 0.7% Score 3; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 270 EER 272

DB 2 EER 4

RESULT 36
 S29637
 lectin beta II chain Actocarpus champeden (fragment)
 C:Species: Actocarpus champeden
 C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1999 #text_change 24-Jul-1998
 C:Accession: S29637
 F:Rago, L.B.; Brillard, M.; Hocke, J.
 Biochim. Biophys. Acta 1156: 215-222, 1993
 A:Title: The alpha and beta subunits of the lectins are cleavage products from a 17
 A:Reference number: S29635, M01D:93152601
 A:Accession: S29637
 A:Molecule type: protein
 A:Residues: 1-6 <NGO>
 A:Experimental source: seed
 C:Complex: heterotetramer; two alpha and two beta chains
 C:Function:
 A:Description: seed storage protein
 A:Note: lectin for D-galactosyl beta-1,3-N-acetylglactosamine
 C:Keywords: heterotetramer; lectin; seed; storage protein

Query Match 0.7% Score 3; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 380 SGK 482

DB 4 SGK 6

RESULT 37
 A60494
 antineoplastic glycoprotein - sea hare (Dolabella auricularia) (fragment)
 N:Alternative names: dolabellin C
 C:Species: Dolabella auricularia
 C:Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 18-Jun-1994
 C:Accession: A60494
 F:Kisaku, J.; Kashi, H.; Yamashita, H.
 Dev. Comp. Immunol. 13: 3-8, 1989
 A:Title: Purification of dolabellin-C an antineoplastic glycoprotein in the body fl
 A:Reference number: A60494, M01D:89457188
 A:Accession: A60494
 A:Molecule type: protein
 A:Residues: 1-6 <KIS>

Accession: EMBL100000; N158497a2.1; E11N.AAB.047.1; E115642a2

Query: B123
 Best Local Similarity: 100.00, Prod. No. 2.06e-05,
 Matches: 0, Conserved: 0, Mismatches: 0, Gaps: 0

27 0 VYS 0
 1 1
 14 1 VYS 4

PSM1 17

Accession:

Query: B123, EMBL100000, Gapwise (Multiple expansion)

Accession: M158497a2.1

Query: B123, EMBL100000, Prod. No. 2.06e-05,
 Matches: 0, Conserved: 0, Mismatches: 0, Gaps: 0

27 0 VYS 0

1 1

14 1 VYS 4

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Accession: M158497a2.1

Job 1-1
3 EXP 3

RESULT 6

FARP_MONEX
ID FARP_MONEX STANDARD PRI 6 AA.

AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FARPamide (FARP) is a peptide that is a sheep anti-FARP.
OS Monodelphis domestica (Sheep tapeworm)
OC Eukaryota; Metazoa; Platyhelminthes; Tricladaria; Platyhelminthes;
OC Rhabdionophora; Euteleostomi; Tricladaria; Platyhelminthes; Tricladaria;
OC Neodermata; Cestoda; Eucestoda; Cyclophyllidae; Anoplocephalidae;
OC Rhabdionophora;
OX NCBI_TaxID: 9941.
FN [1]
FN SEQUENCE
EX MEDLINE: 93312289; PubMed: 8424541;
FA Magro A.G., Shaw C., Hinton D.W., Kim L.,
FI "GNFPPamide: a novel FARPamide-immunoreactive peptide isolated from
FI the sheep tapeworm, Monodelphis domestica."
FI Biochem. Biophys. Res. Commun. 193:1054-1060(1993).
CC -1- SIMILARITY: BELONGS TO THE FARP (FARPAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FI MEDRES 6 6 AMIDATION.
SQ SEQUENCE 6 AA: 787 MW: 6940969C4481000 CRG64.

Query Match 0.7%; Score 3; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. le+05; Mismatches 0; Indels 0; Caps 0;

CY 2-3 GNF 239

DB 1 GNF 3

RESULT 7

FARP_MONEX
ID FARP_MONEX STANDARD PRI 6 AA.

AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Locustastatin I.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Phrynosoma; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID: 7004;
FN [1]
FN SEQUENCE
EX MEDLINE: 92262461; PubMed: 1565017;
FA Schools L., Holman G.M., Probst P., van Damme J., Hayes T.K.,
FA de Laet A.,
FI "Locustastatin, a novel myotropic peptide from Locusta migratoria,
FI isolation, primary structure and synthesis."
FI Regul. Pept. 47:49-57(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION
CC OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN
CC TUBULES.
DR PIR: A61058; A61058.
KW Neuropeptide; Amidation.
FI MEDRES 6 6 AMIDATION.
SQ SEQUENCE 6 AA: 654 MW: 686465A5B034000 CRG64.

Query Match 0.7%; Score 3; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. le+05; Mismatches 0; Indels 0; Caps 0;

CY 3-3 SWG 365

DB 4 SWG 6

RESULT 8

OVMLLEPDE
ID OVMLLEPDE STANDARD PRI 6 AA.

AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Oviductal motility stimulating peptide (leb-ovm).
OS Lepidoptera; decaimata (Colorado potato beetle).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Chrysomelidae; Chrysomelidae; Chrysomelidae;
OX NCBI_TaxID: 7549;
FN [1]
FN SEQUENCE AND SYNTHESIS.
EX TISSUE: Head;
FA MEDLINE: 91271880; PubMed: 2052497;
FA Spittael K., Schoofs L., Grauwels L., Smet H., van Damme J.,
FA Probst P., Torrekens S., de Loof A.,
FI "Isolation, identification and synthesis of novel oviductal motility
FI stimulating head peptide in the Colorado potato beetle, Leptinotarsa
FI decemlineata."
FI Regul. Peptides 12:31-36(1991).
CC -1- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
CC OVIDUCT.
KW Neuropeptide; Amidation.
FI MEDRES 6 6 AMIDATION.
SQ SEQUENCE 6 AA: 720 MW: 6B07642B5DD03000 CRG64.

Query Match 0.7%; Score 3; DB 1; Length 6;

Best Local Similarity 100.0%; Pred. No. le+05; Mismatches 0; Indels 0; Caps 0;

CY 289 AYK 291

DB 2 AYK 4

RESULT 9

GRWM_HUMAN
ID GRWM_HUMAN STANDARD PRI 3 AA.

AC 01-NOV-1995 (Rel. 32, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID: 9606;
FN [1]
FN SEQUENCE.
EX MEDLINE: 77452469; PubMed: 958356;
FA Schlesinger D.H., Pickart L., Thaler M.,
FI "Growth-modulating serum tripeptide is glycyl-histidyl-lysine."
FI Regul. Pept. 33:323-325(1977).
CC -1- MISCELLANEOUS: THIS SEQUENCE TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
DR PIR: A01421; GKHU.
SQ SEQUENCE 3 AA: 340 MW: 6341E1000000000 CRG64.

Query Match 0.5%; Score 2; DB 1; Length 3;

Best Local Similarity 100.0%; Pred. No. le+05; Mismatches 0; Indels 0; Caps 0;

CY 94 HK 95

Age Group	1990	2000	2010	2020
0-14	15%	12%	10%	8%
15-24	12%	10%	8%	6%
25-34	10%	8%	6%	4%
35-44	8%	6%	4%	2%
45-54	6%	4%	2%	1%
55-64	4%	2%	1%	0%
65-74	2%	1%	0%	0%
75+	1%	0%	0%	0%

TABLE 2. *SWISS* plot and χ^2 test results. The χ^2 plot used the single a correlation coefficient between the Swiss and Dutch two informants, and the χ^2 test used the χ^2 test between the Swiss and Dutch two informants. There are two possible outcomes for the χ^2 test: significant (sig.) and not significant (n.s.). The χ^2 test is based on the χ^2 test for independence.


```

KA Plant Cell Physiol. 41:27-32(2000).
CC -1- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
CC CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
CC EMBRYOS.
CC -1- SUPPLEMENTAR LOCATION: SECRETED.
CC -1- TIME REGULATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
CC PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
KW GROWTH FACTOR: Sulfation.
FT PEPTIDE 1 4 PHYTOSULFOKINE-HEXA.
FT MODRES 1 1 Sulfation.
FT MODRES 3 3 Sulfation.
SQ SEQUENCE 5 AA: 667 MW: 760188504360600 CRC64;

Query Match 0.5%; Score 2; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 YI 474
I
I YI 2

RESULT 34
RE32_LITRU
ID RE32_LITRU STANDARD: PRI; 5 AA.
AC P82070;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Rubellidin 1.1.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Hyliidae;
CC Litoria.
CC RCP: TaxID=104895.
CC [1]
CC SEQUENCE, AND MASS SPECTROMETRY.
CC TISSUE SKIN SECRETION.
KA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Howie J.H., Gao C.,
KA Tyler M.J., Wallace J.C.;
PT "The structure of new peptides from the Australian red tree frog
PT 'Litoria rubella', the skin peptide profile as a probe for the study
PT of evolutionary trends of amphibians."
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIHISTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=548; METHOD=FAE.
KW Amphibian skin.
SQ SEQUENCE 5 AA: 598 MW: 600909048200000 CRC64;

Query Match 0.5%; Score 2; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 YI 241
I
I YI 2

RESULT 35
RE32_LITRU
ID RE32_LITRU STANDARD: PRI; 5 AA.
AC P82071;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Rubellidin 2.1.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Hyliidae;
CC Litoria.

```

```

OX NCBI_TaxID=104895;
KA [1]
CC SEQUENCE, AND MASS SPECTROMETRY.
CC TISSUE SKIN SECRETION.
KA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Howie J.H., Gao C.,
KA Tyler M.J., Wallace J.C.;
PT "The structure of new peptides from the Australian red tree frog
PT 'Litoria rubella', the skin peptide profile as a probe for the study
PT of evolutionary trends of amphibians."
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIHISTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=626; METHOD=FAE.
KW Amphibian skin.
SQ SEQUENCE 5 AA: 626 MW: 600909048200000 CRC64;

Query Match 0.5%; Score 2; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 YI 284
I
I YI 2

RESULT 36
RE32_LITRU
ID RE32_LITRU STANDARD: PRI; 5 AA.
AC P82072;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Rubellidin 3.1.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Hyliidae;
CC Litoria.
CC NCBI_TaxID=104895;
CC [1]
CC SEQUENCE, AND MASS SPECTROMETRY.
CC TISSUE SKIN SECRETION.
KA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Howie J.H., Gao C.,
KA Tyler M.J., Wallace J.C.;
PT "The structure of new peptides from the Australian red tree frog
PT 'Litoria rubella', the skin peptide profile as a probe for the study
PT of evolutionary trends of amphibians."
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIHISTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAE.
KW Amphibian skin; Amidation.
FT MODRES 5 5 AMIDATION.
SQ SEQUENCE 5 AA: 655 MW: 716079010200000 CRC64;

Query Match 0.5%; Score 2; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 YI 284
I
I YI 2

RESULT 37
RE32_LITRU
ID RE32_LITRU STANDARD: PRI; 5 AA.
AC P82073;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Rubellidin 3.2.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: october 9, 2002, 11:51:36 ; Search time 29 seconds
(without alignments)
2392.102 Million cell updates/sec

Titles: US-09-598-062-4
RefSeq score: 40;
Sequence: 1 MLIUNNVEERQEVYISPYIA.....KSHICGLLVEPVVPIINQSI 401

Scoring table: 0.000
Gapop 60.0 , Gapext 60.0
Searched: 562222 seqs, 17299429 residues
Word size : 0

Total number of hits satisfying chosen parameters: 13

Minimum (as seq length): 0
Maximum (as seq length): 6

Post-processing: listing first 1000 summaries

Database : SPTRMBL196*
1: SP archaea*
2: SP bacteria*
3: SP fungi*
4: SP human*
5: SP invertebrate*
6: SP mammal*
7: SP mhc*
8: SP organelle*
9: SP plaque*
10: SP plant*
11: SP rodent*
12: SP virus*
13: SP vertebrate*
14: SP unclassified*
15: SP rivirus*
16: SP bacteriopl*
17: SP archaeopl*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3	0.7	6	13	P82096 Litoria rub
2	2	0.5	4	11	Q08433 Rattus norv
3	2	0.5	5	2	P63073 Lactillus ce
4	2	0.5	5	10	Q59007 Rattus norv
5	2	0.5	4	13	P82070 Litoria rub
6	2	0.5	4	13	P82071 Litoria rub
7	2	0.5	5	13	P82072 Litoria rub
8	2	0.5	5	13	P82073 Litoria rub
9	2	0.5	5	13	P82099 Litoria rub
10	2	0.5	5	13	P82100 Litoria rub
11	2	0.5	4	10	P82101 Spinaelia ol
12	2	0.5	6	10	P82541 Spinaelia ol
13	2	0.5	6	10	P82192 Spinaelia ol

ALIGNMENTS

RESULT 1
P82096
ID P82096 PRELIMINARY: PRT: 6 AA.
AC P82096
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
RT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ELECTRIN 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID:104845;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN SECRETION;
EA Wabell, P.A., Baxie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:0-0(1999).
KW Amphibian skin; Amidation.
FT MOD_RES 6
SQ SEQUENCE 5 AA; 792 MW; 64879477299AA090 CRC64;
AMIDATION.
Query Match 0.7%; Score 3; DB 13; Length 6;
Best Local Similarity 100.0%; Prod. No. 5,6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 FVP 71
DB 1 FVP 3
RESULT 2
Q08433
ID Q08433 PRELIMINARY: PRT: 4 AA.
AC Q08433
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DI 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DI UDP-GLUTAMINOSYLTRANSFERASE, MICROSMAL (EC 2.4.1.17) (UDPQT)
DE (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUNN;
RX MEDLINE:91282758; PubMed 1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DE FUNCTION, UDPQT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
DE SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
DE SUBSEQUENT COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR UDP - ACCEPTOR
CC -1- BETA-D-GLUCURONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROsome.
DR EMBL: S48636; AAB19259.1; -;
KW transferase, glycosyltransferase; Microsome; Multi gene family.
FT NON_TER 1 1
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 473 MW; 647792642000000000 CRC64;
Query Match 0.5%; Score 2; DB 11; Length 4;
Best Local Similarity 100.0%; Prod. No. 5,6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Figure 1 consists of two parts, (a) and (b), illustrating the experimental design. Part (a) is labeled 'Pretest' and shows a subject in a car, looking at a screen. The screen displays a target (T) and a starting point (S). The car is moving towards the target. Part (b) is labeled 'Main experiment' and shows a subject in a car, looking at a screen. The screen displays a target (T) and a starting point (S). The car is moving towards the target. The screen also displays a scale from 0 to 100 cm.

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

<

100

1. $\frac{1}{2}$ 2. $\frac{1}{3}$ 3. $\frac{1}{4}$ 4. $\frac{1}{5}$ 5. $\frac{1}{6}$ 6. $\frac{1}{7}$ 7. $\frac{1}{8}$ 8. $\frac{1}{9}$ 9. $\frac{1}{10}$ 10. $\frac{1}{11}$ 11. $\frac{1}{12}$ 12. $\frac{1}{13}$ 13. $\frac{1}{14}$ 14. $\frac{1}{15}$ 15. $\frac{1}{16}$ 16. $\frac{1}{17}$ 17. $\frac{1}{18}$ 18. $\frac{1}{19}$ 19. $\frac{1}{20}$ 20. $\frac{1}{21}$ 21. $\frac{1}{22}$ 22. $\frac{1}{23}$ 23. $\frac{1}{24}$ 24. $\frac{1}{25}$ 25. $\frac{1}{26}$ 26. $\frac{1}{27}$ 27. $\frac{1}{28}$ 28. $\frac{1}{29}$ 29. $\frac{1}{30}$ 30. $\frac{1}{31}$ 31. $\frac{1}{32}$ 32. $\frac{1}{33}$ 33. $\frac{1}{34}$ 34. $\frac{1}{35}$ 35. $\frac{1}{36}$ 36. $\frac{1}{37}$ 37. $\frac{1}{38}$ 38. $\frac{1}{39}$ 39. $\frac{1}{40}$ 40. $\frac{1}{41}$ 41. $\frac{1}{42}$ 42. $\frac{1}{43}$ 43. $\frac{1}{44}$ 44. $\frac{1}{45}$ 45. $\frac{1}{46}$ 46. $\frac{1}{47}$ 47. $\frac{1}{48}$ 48. $\frac{1}{49}$ 49. $\frac{1}{50}$ 50. $\frac{1}{51}$ 51. $\frac{1}{52}$ 52. $\frac{1}{53}$ 53. $\frac{1}{54}$ 54. $\frac{1}{55}$ 55. $\frac{1}{56}$ 56. $\frac{1}{57}$ 57. $\frac{1}{58}$ 58. $\frac{1}{59}$ 59. $\frac{1}{60}$ 60. $\frac{1}{61}$ 61. $\frac{1}{62}$ 62. $\frac{1}{63}$ 63. $\frac{1}{64}$ 64. $\frac{1}{65}$ 65. $\frac{1}{66}$ 66. $\frac{1}{67}$ 67. $\frac{1}{68}$ 68. $\frac{1}{69}$ 69. $\frac{1}{70}$ 70. $\frac{1}{71}$ 71. $\frac{1}{72}$ 72. $\frac{1}{73}$ 73. $\frac{1}{74}$ 74. $\frac{1}{75}$ 75. $\frac{1}{76}$ 76. $\frac{1}{77}$ 77. $\frac{1}{78}$ 78. $\frac{1}{79}$ 79. $\frac{1}{80}$ 80. $\frac{1}{81}$ 81. $\frac{1}{82}$ 82. $\frac{1}{83}$ 83. $\frac{1}{84}$ 84. $\frac{1}{85}$ 85. $\frac{1}{86}$ 86. $\frac{1}{87}$ 87. $\frac{1}{88}$ 88. $\frac{1}{89}$ 89. $\frac{1}{90}$ 90. $\frac{1}{91}$ 91. $\frac{1}{92}$ 92. $\frac{1}{93}$ 93. $\frac{1}{94}$ 94. $\frac{1}{95}$ 95. $\frac{1}{96}$ 96. $\frac{1}{97}$ 97. $\frac{1}{98}$ 98. $\frac{1}{99}$ 99. $\frac{1}{100}$ 100. $\frac{1}{101}$ 101. $\frac{1}{102}$ 102. $\frac{1}{103}$ 103. $\frac{1}{104}$ 104. $\frac{1}{105}$ 105. $\frac{1}{106}$ 106. $\frac{1}{107}$ 107. $\frac{1}{108}$ 108. $\frac{1}{109}$ 109. $\frac{1}{110}$ 110. $\frac{1}{111}$ 111. $\frac{1}{112}$ 112. $\frac{1}{113}$ 113. $\frac{1}{114}$ 114. $\frac{1}{115}$ 115. $\frac{1}{116}$ 116. $\frac{1}{117}$ 117. $\frac{1}{118}$ 118. $\frac{1}{119}$ 119. $\frac{1}{120}$ 120. $\frac{1}{121}$ 121. $\frac{1}{122}$ 122. $\frac{1}{123}$ 123. $\frac{1}{124}$ 124. $\frac{1}{125}$ 125. $\frac{1}{126}$ 126. $\frac{1}{127}$ 127. $\frac{1}{128}$ 128. $\frac{1}{129}$ 129. $\frac{1}{130}$ 130. $\frac{1}{131}$ 131. $\frac{1}{132}$ 132. $\frac{1}{133}$ 133. $\frac{1}{134}$ 134. $\frac{1}{135}$ 135. $\frac{1}{136}$ 136. $\frac{1}{137}$ 137. $\frac{1}{138}$ 138. $\frac{1}{139}$ 139. $\frac{1}{140}$ 140. $\frac{1}{141}$ 141. $\frac{1}{142}$ 142. $\frac{1}{143}$ 143. $\frac{1}{144}$ 144. $\frac{1}{145}$ 145. $\frac{1}{146}$ 146. $\frac{1}{147}$ 147. $\frac{1}{148}$ 148. $\frac{1}{149}$ 149. $\frac{1}{150}$ 150. $\frac{1}{151}$ 151. $\frac{1}{152}$ 152. $\frac{1}{153}$ 153. $\frac{1}{154}$ 154. $\frac{1}{155}$ 155. $\frac{1}{156}$ 156. $\frac{1}{157}$ 157. $\frac{1}{158}$ 158. $\frac{1}{159}$ 159. $\frac{1}{160}$ 160. $\frac{1}{161}$ 161. $\frac{1}{162}$ 162. $\frac{1}{163}$ 163. $\frac{1}{164}$ 164. $\frac{1}{165}$ 165. $\frac{1}{166}$ 166. $\frac{1}{167}$ 167. $\frac{1}{168}$ 168. $\frac{1}{169}$ 169. $\frac{1}{170}$ 170. $\frac{1}{171}$ 171. $\frac{1}{172}$ 172. $\frac{1}{173}$ 173. $\frac{1}{174}$ 174. $\frac{1}{175}$ 175. $\frac{1}{176}$ 176. $\frac{1}{177}$ 177. $\frac{1}{178}$ 178. $\frac{1}{179}$ 179. $\frac{1}{180}$ 180. $\frac{1}{181}$ 181. $\frac{1}{182}$ 182. $\frac{1}{183}$ 183. $\frac{1}{184}$ 184. $\frac{1}{185}$ 185. $\frac{1}{186}$ 186. $\frac{1}{187}$ 187. $\frac{1}{188}$ 188. $\frac{1}{189}$ 189. $\frac{1}{190}$ 190. $\frac{1}{191}$ 191. $\frac{1}{192}$ 192. $\frac{1}{193}$ 193. $\frac{1}{194}$ 194. $\frac{1}{195}$ 195. $\frac{1}{196}$ 196. $\frac{1}{197}$ 197. $\frac{1}{198}$ 198. $\frac{1}{199}$ 199. $\frac{1}{200}$ 200. $\frac{1}{201}$ 201. $\frac{1}{202}$ 202. $\frac{1}{203}$ 203. $\frac{1}{204}$ 204. $\frac{1}{205}$ 205. $\frac{1}{206}$ 206. $\frac{1}{207}$ 207. $\frac{1}{208}$ 208. $\frac{1}{209}$ 209. $\frac{1}{210}$ 210. $\frac{1}{211}$ 211. $\frac{1}{212}$ 212. $\frac{1}{213}$ 213. $\frac{1}{214}$ 214. $\frac{1}{215}$ 215. $\frac{1}{216}$ 216. $\frac{1}{217}$ 217. $\frac{1}{218}$ 218. $\frac{1}{219}$ 219. $\frac{1}{220}$ 220. $\frac{1}{221}$ 221. $\frac{1}{222}$ 222. $\frac{1}{223}$ 223. $\frac{1}{224}$ 224. $\frac{1}{225}$ 225. $\frac{1}{226}$ 226. $\frac{1}{227}$ 227. $\frac{1}{228}$ 228. $\frac{1}{229}$ 229. $\frac{1}{230}$ 230. $\frac{1}{231}$ 231. $\frac{1}{232}$ 232. $\frac{1}{233}$ 233. $\frac{1}{234}$ 234. $\frac{1}{235}$ 235. $\frac{1}{236}$ 236. $\frac{1}{237}$ 237. $\frac{1}{238}$ 238. $\frac{1}{239}$ 239. $\frac{1}{240}$ 240.

100

10

7

$\frac{1}{n} \sum_{i=1}^n \log p_i$

100

2. 2. 2.

10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

100

10

100

100

100

2 3 4

10

1.
 2.
 3.
 4.
 5.
 6.
 7.
 8.
 9.
 10.
 11.
 12.
 13.
 14.
 15.
 16.
 17.
 18.
 19.
 20.
 21.
 22.
 23.
 24.
 25.
 26.
 27.
 28.
 29.
 30.

1	2
3	4
5	6
7	8
9	10
11	12
13	14
15	16
17	18
19	20
21	22
23	24
25	26
27	28
29	30
31	32
33	34
35	36
37	38
39	40
41	42
43	44
45	46
47	48
49	50
51	52
53	54
55	56
57	58
59	60
61	62
63	64
65	66
67	68
69	70
71	72
73	74
75	76
77	78
79	80
81	82
83	84
85	86
87	88
89	90
91	92
93	94
95	96
97	98
99	100

Figure 1

圖 10-10 鋼筋的斷面

Figure 1. Schematic representation of the experimental design. The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG).

[illegible]

1. **Introduction**
 2. **Background**
 3. **Methodology**
 4. **Results**
 5. **Conclusion**
 6. **References**
 7. **Appendix**
 8. **Figure 1**
 9. **Figure 2**
 10. **Figure 3**
 11. **Figure 4**
 12. **Figure 5**
 13. **Figure 6**
 14. **Figure 7**
 15. **Figure 8**
 16. **Figure 9**
 17. **Figure 10**
 18. **Figure 11**
 19. **Figure 12**
 20. **Figure 13**
 21. **Figure 14**
 22. **Figure 15**
 23. **Figure 16**
 24. **Figure 17**
 25. **Figure 18**
 26. **Figure 19**
 27. **Figure 20**
 28. **Figure 21**
 29. **Figure 22**
 30. **Figure 23**
 31. **Figure 24**
 32. **Figure 25**
 33. **Figure 26**
 34. **Figure 27**
 35. **Figure 28**
 36. **Figure 29**
 37. **Figure 30**
 38. **Figure 31**
 39. **Figure 32**
 40. **Figure 33**
 41. **Figure 34**
 42. **Figure 35**
 43. **Figure 36**
 44. **Figure 37**
 45. **Figure 38**
 46. **Figure 39**
 47. **Figure 40**
 48. **Figure 41**
 49. **Figure 42**
 50. **Figure 43**
 51. **Figure 44**
 52. **Figure 45**
 53. **Figure 46**
 54. **Figure 47**
 55. **Figure 48**
 56. **Figure 49**
 57. **Figure 50**
 58. **Figure 51**
 59. **Figure 52**
 60. **Figure 53**
 61. **Figure 54**
 62. **Figure 55**
 63. **Figure 56**
 64. **Figure 57**
 65. **Figure 58**
 66. **Figure 59**
 67. **Figure 60**
 68. **Figure 61**
 69. **Figure 62**
 70. **Figure 63**
 71. **Figure 64**
 72. **Figure 65**
 73. **Figure 66**
 74. **Figure 67**
 75. **Figure 68**
 76. **Figure 69**
 77. **Figure 70**
 78. **Figure 71**
 79. **Figure 72**
 80. **Figure 73**
 81. **Figure 74**
 82. **Figure 75**
 83. **Figure 76**
 84. **Figure 77**
 85. **Figure 78**
 86. **Figure 79**
 87. **Figure 80**
 88. **Figure 81**
 89. **Figure 82**
 90. **Figure 83**
 91. **Figure 84**
 92. **Figure 85**
 93. **Figure 86**
 94. **Figure 87**
 95. **Figure 88**
 96. **Figure 89**
 97. **Figure 90**
 98. **Figure 91**
 99. **Figure 92**
 100. **Figure 93**
 101. **Figure 94**
 102. **Figure 95**
 103. **Figure 96**
 104. **Figure 97**
 105. **Figure 98**
 106. **Figure 99**
 107. **Figure 100**
 108. **Figure 101**
 109. **Figure 102**
 110. **Figure 103**
 111. **Figure 104**
 112. **Figure 105**
 113. **Figure 106**
 114. **Figure 107**
 115. **Figure 108**
 116. **Figure 109**
 117. **Figure 110**
 118. **Figure 111**
 119. **Figure 112**
 120. **Figure 113**
 121. **Figure 114**
 122. **Figure 115**
 123. **Figure 116**
 124. **Figure 117**
 125. **Figure 118**
 126. **Figure 119**
 127. **Figure 120**
 128. **Figure 121**
 129. **Figure 122**
 130. **Figure 123**
 131. **Figure 124**
 132. **Figure 125**
 133. **Figure 126**
 134. **Figure 127**
 135. **Figure 128**
 136. **Figure 129**
 137. **Figure 130**
 138. **Figure 131**
 139. **Figure 132**
 140. **Figure 133**
 141. **Figure 134**
 142. **Figure 135**
 143. **Figure 136**
 144. **Figure 137**
 145. **Figure 138**
 146. **Figure 139**
 147. **Figure 140**
 148. **Figure 141**
 149. **Figure 142**
 150. **Figure 143**
 151. **Figure 144**
 152. **Figure 145**
 153. **Figure 146**
 154. **Figure 147**
 155. **Figure 148**
 156. **Figure 149**
 157. **Figure 150**
 158. **Figure 151**
 159. **Figure 152**
 160. **Figure 153**
 161. **Figure 154**
 162. **Figure 155**
 163. **Figure 156**
 164. **Figure 157**
 165. **Figure 158**
 166. **Figure 159**
 167. **Figure 160**
 168. **Figure 161**
 169. **Figure 162**
 170. **Figure 163**
 171. **Figure 164**
 172. **Figure 165**
 173. **Figure 166**
 174. **Figure 167**
 175. **Figure 168**
 176. **Figure 169**
 177. **Figure 170**
 178. **Figure 171**
 179. **Figure 172**
 180. **Figure 173**
 181. **Figure 174**
 182. **Figure 175**
 183. **Figure 176**
 184. **Figure 177**
 185. **Figure 178**
 186. **Figure 179**
 187. **Figure 180**
 188. **Figure 181**
 189. **Figure 182**
 190. **Figure 183**
 191. **Figure 184**
 192. **Figure 185**
 193. **Figure 186**
 194. **Figure 187**
 195. **Figure 188**
 196. **Figure 189**
 197. **Figure 190**
 198. **Figure 191**
 199. **Figure 192**
 200. **Figure 193**
 201. **Figure 194**
 202. **Figure 195**
 203. **Figure 196**
 204. **Figure 197**
 205. **Figure 198**
 206. **Figure 199**
 207. **Figure 200**
 208. **Figure 201**
 209. **Figure 202**
 210. **Figure 203**
 211. **Figure 204**
 212. **Figure 205**
 213. **Figure 206**
 214. **Figure 207**
 215. **Figure 208**
 216. **Figure 209**
 217. **Figure 210</**

1.
 2.
 3.
 4.
 5.
 6.
 7.
 8.
 9.
 10.
 11.
 12.
 13.
 14.
 15.
 16.
 17.
 18.
 19.
 20.
 21.
 22.
 23.
 24.
 25.
 26.
 27.
 28.
 29.
 30.

Figure 1. Schematic representation of the experimental design. The subjects were divided into two groups: the control group (C) and the experimental group (E). The control group (C) was divided into two subgroups: the control group (C) and the control group (C). The experimental group (E) was divided into two subgroups: the experimental group (E) and the experimental group (E). The control group (C) was divided into two subgroups: the control group (C) and the control group (C). The experimental group (E) was divided into two subgroups: the experimental group (E) and the experimental group (E).

$$\begin{array}{c} \mathbb{Z} \\ \mathbb{Z} \\ \mathbb{Z} \end{array}$$

The figure consists of two bar charts side-by-side. The left chart shows vaccination percentages for respondents aged 18-29, and the right chart shows percentages for respondents aged 30-39. Both charts compare vaccination status across five categories: Not vaccinated at all, Vaccinated once, Vaccinated twice, Vaccinated three times, and Vaccinated four or more times.

Vaccination Status	18-29 Age Group (%)	30-39 Age Group (%)
Not vaccinated at all	~7%	~1%
Vaccinated once	~15%	~1%
Vaccinated twice	~25%	~1%
Vaccinated three times	~35%	~1%
Vaccinated four or more times	~20%	~1%

1

1. $\frac{1}{2}$
 2. $\frac{1}{2}$
 3. $\frac{1}{2}$
 4. $\frac{1}{2}$
 5. $\frac{1}{2}$
 6. $\frac{1}{2}$
 7. $\frac{1}{2}$
 8. $\frac{1}{2}$
 9. $\frac{1}{2}$
 10. $\frac{1}{2}$

•

1

Matches 2: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 QY 283 1E 284
 DB 1 1E 2

RESULT 7

PR2072 PRELIMINARY: PRT: 5 AA.
 AC PR2072:
 DT 01-MAY-2000 (TREMURel. 13, created)
 DT 01-MAY-2000 (TREMURel. 13, last sequence update)
 DE 01-MAY-2000 (TREMURel. 13, last annotation update)
 DE RUBELLIDIN 3.1.
 OS Litoria rubella (Desert tree frog).
 OS Fukuyuta; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OC NCBI_TaxID=104895;

KN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE-SKIN SECRETION.
 RA Steinhorner S.T., Wabnitz P.A., Waugh P.J., Howie J.H., Gao C.,
 RA Tyler M.J., Wallace J.C.;
 RT 'The structure of new peptides from the Australian red tree frog
 RT 'Litoria rubella', the skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians.';
 RL Aust. J. Chem. 49:955-964(1996).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIBIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SPECTRED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAE.
 CC Amphibian skin; Amidation.
 FT MOD_RES 5
 FT SEQUENCE 5 AA; 656 MW; 71A9C9C810300000 CRC64;

Query Match 0.5%: Score 2: DB 13: Length 5;
 Best Local Similarity 100.0%: Fred. No. 5.6e-05;
 Matches 2: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 283 1E 284
 DB 1 1E 2

RESULT 8

PR2073 PRELIMINARY: PRT: 5 AA.
 AC PR2073:
 DT 01-MAY-2000 (TREMURel. 13, created)
 DT 01-MAY-2000 (TREMURel. 13, last sequence update)
 DT 01-MAY-2000 (TREMURel. 13, last annotation update)
 DE RUBELLIDIN 3.2.
 OS Litoria rubella (Desert tree frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OC NCBI_TaxID=104895;

KN [1]
 RP SEQUENCE.
 RC TISSUE-SKIN SECRETION.
 RA Wabnitz P.A., Howie J.H., Tyler M.J., Wallace J.C.;
 RT 'Peptides from the skin glands of the Australian buzzing tree frog
 RT 'Litoria rubella', comparison with the skin peptides from Litoria
 RT 'rubella'.';
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIBIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC Amphibian skin.
 FT SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 0.5%: Score 2: DB 13: Length 5;
 Best Local Similarity 100.0%: Fred. No. 5.6e-05;
 Matches 2: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 124 1E 125
 DB 2 1E 3

RESULT 9

PR2099 PRELIMINARY: PRT: 5 AA.
 AC PR2099:
 DT 01-MAY-2000 (TREMURel. 13, created)
 DT 01-MAY-2000 (TREMURel. 13, last sequence update)
 DE 01-MAY-2000 (TREMURel. 13, last annotation update)
 DE ELECTRIN 3.
 OS Litoria rubella (Desert tree frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OC NCBI_TaxID=104895;

KN [1]
 RP SEQUENCE.
 RC TISSUE-SKIN SECRETION.
 RA Wabnitz P.A., Howie J.H., Tyler M.J., Wallace J.C.;
 RT 'Peptides from the skin glands of the Australian buzzing tree frog
 RT 'Litoria rubella', comparison with the skin peptides from Litoria
 RT 'rubella'.';
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIBIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SPECTRED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAE.
 CC Amphibian skin; Amidation.
 FT MOD_RES 5
 FT SEQUENCE 5 AA; 660 MW; 6697A1F2C9A00000 CRC64;

Query Match 0.5%: Score 2: DB 13: Length 5;
 Best Local Similarity 100.0%: Fred. No. 5.6e-05;
 Matches 2: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 229 1E 230
 DB 1 1E 2

RESULT 10

PR2100 PRELIMINARY: PRT: 5 AA.
 AC PR2100:
 DT 01-MAY-2000 (TREMURel. 13, created)
 DT 01-MAY-2000 (TREMURel. 13, last sequence update)
 DT 01-MAY-2000 (TREMURel. 13, last annotation update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OC NCBI_TaxID=104895;

KN [1]
 RP SEQUENCE.
 RC TISSUE-SKIN SECRETION.
 RA Wabnitz P.A., Howie J.H., Tyler M.J., Wallace J.C.;
 RT 'Peptides from the skin glands of the Australian buzzing tree frog
 RT 'Litoria rubella', comparison with the skin peptides from Litoria
 RT 'rubella'.';
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIBIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC Amphibian skin; Amidation.
 FT MOD_RES 5
 FT SEQUENCE 5 AA; 616 MW; 61E2D1AD05A00000 CRC64;

Query Match 0.5%: Score 2: DB 13: Length 5;
 Best Local Similarity 100.0%: Fred. No. 5.6e-05;
 Matches 2: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 47 1E 48

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

GM protein - protein search, using sw model

Run on: October 9, 2002, 11:57:11 : Search time 19 Seconds

(without alignments)
515,509 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 401

Sequence: 1 MGLGNVEEHCVISQVYA.....KRGVGLVMEVYVINGSI 401

Scoring table: 0.100

Gapop 60.0 : Gapext 60.0

Searched: 231628 scds, 24125734 residues

Word size : 0

Total number of hits satisfying chosen parameters: 20957

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: listing first 1000 summaries

Database : Issued patents AA:*

1: /pat26/patent/12/1047A_Comp pep.*
2: /pat26/patent/12/1047A_Comp pep.*
3: /pat26/patent/12/1047A_Comp pep.*
4: /pat26/patent/12/1047A_Comp pep.*
5: /pat26/patent/12/1047A_Comp pep.*
6: /pat26/patent/12/1047A_Comp pep.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No	Score	% Match	Length	DB	ID	Description
1	5	1.2	5	4	US-08-827-171B-15	Sequence 15, App1
2	5	1.2	6	1	US-07-752-101A-5	Sequence 9, App1
3	5	1.2	6	1	US-08-093-741-72	Sequence 72, App1
4	5	1.2	6	1	US-08-486-721A-19	Sequence 19, App1
5	5	1.2	6	1	US-08-720-012-72	Sequence 72, App1
6	5	1.2	6	2	US-07-662-741D-41	Sequence 41, App1
7	4	1.0	4	1	US-08-361-862-4	Sequence 4, App1
8	4	1.0	4	2	US-08-623-833B-67	Sequence 67, App1
9	4	1.0	4	2	US-08-392-328A-7	Sequence 7, App1
10	4	1.0	4	2	US-08-332-539D-52	Sequence 52, App1
11	4	1.0	4	3	US-09-166-750-7	Sequence 7, App1
12	4	1.0	4	3	US-09-166-093-7	Sequence 7, App1
13	4	1.0	4	3	US-09-172-019-7	Sequence 7, App1
14	4	1.0	4	3	US-09-166-054-7	Sequence 7, App1
15	4	1.0	4	4	US-09-236-160-54	Sequence 54, App1
16	4	1.0	4	4	5521089-9	Patent No.5521089
17	4	1.0	5	1	US-08-456-840-34	Sequence 34, App1
18	4	1.0	5	1	US-08-340-428B-46	Sequence 46, App1
19	4	1.0	5	1	US-08-211-717-6	Sequence 6, App1
20	4	1.0	5	1	US-08-196-721A-29	Sequence 20, App1
21	4	1.0	5	1	US-08-220-101-19	Sequence 19, App1
22	4	1.0	5	1	US-08-157-607-10	Sequence 10, App1
23	4	1.0	5	1	US-08-266-407A-34	Sequence 34, App1
24	4	1.0	5	1	US-08-655-220-25	Sequence 25, App1
25	4	1.0	5	1	US-08-127-252-15	Sequence 15, App1
26	4	1.0	5	1	US-08-469-302-10	Sequence 10, App1
27	4	1.0	5	2	US-08-618-108B-25	Sequence 25, App1

28	4	1.0	5	2	US-08-850-392-10	Sequence 10, App1
29	4	1.0	5	2	US-08-892-544-14	Sequence 44, App1
30	4	1.0	5	2	US-08-637-279-31	Sequence 31, App1
31	4	1.0	5	2	US-08-928-958-20	Sequence 20, App1
32	4	1.0	5	2	US-08-742-440A-14	Sequence 14, App1
33	4	1.0	5	2	US-08-450-260A-41B	Sequence 41B, App1
34	4	1.0	5	2	US-08-755-728-6	Sequence 6, App1
35	4	1.0	5	2	US-08-974-655-46	Sequence 46, App1
36	4	1.0	5	2	US-08-313-562A-6	Sequence 6, App1
37	4	1.0	5	4	US-08-871-356A-81	Sequence 81, App1
38	4	1.0	5	4	US-08-144-778-14	Sequence 14, App1
39	4	1.0	5	4	US-09-062-894A-11	Sequence 11, App1
40	4	1.0	5	4	US-09-201-945-81	Sequence 81, App1
41	4	1.0	5	4	US-09-187-859-43A	Sequence 43A, App1
42	4	1.0	5	4	US-09-187-859-125A	Sequence 125A, App1
43	4	1.0	5	4	US-09-187-859-166A	Sequence 166A, App1
44	4	1.0	5	4	US-09-187-859-167A	Sequence 167A, App1
45	4	1.0	5	4	US-09-187-859-169A	Sequence 169A, App1
46	4	1.0	5	4	US-09-187-859-170A	Sequence 170A, App1
47	4	1.0	5	4	US-09-187-859-171A	Sequence 171A, App1
48	4	1.0	5	4	US-09-187-859-172A	Sequence 172A, App1
49	4	1.0	5	4	US-09-187-859-173A	Sequence 173A, App1
50	4	1.0	5	4	US-09-187-859-174A	Sequence 174A, App1
51	4	1.0	5	4	US-09-187-859-175A	Sequence 175A, App1
52	4	1.0	5	4	US-09-187-859-176A	Sequence 176A, App1
53	4	1.0	5	4	US-09-187-859-177A	Sequence 177A, App1
54	4	1.0	5	4	US-09-187-859-178A	Sequence 178A, App1
55	4	1.0	5	4	US-09-187-859-179A	Sequence 179A, App1
56	4	1.0	5	4	US-09-187-859-180A	Sequence 180A, App1
57	4	1.0	5	4	US-09-187-859-181A	Sequence 181A, App1
58	4	1.0	5	4	US-09-187-859-182A	Sequence 182A, App1
59	4	1.0	5	4	US-09-187-859-183A	Sequence 183A, App1
60	4	1.0	5	4	US-09-187-859-184A	Sequence 184A, App1
61	4	1.0	5	4	US-09-187-859-185A	Sequence 185A, App1
62	4	1.0	5	4	US-09-187-859-186A	Sequence 186A, App1
63	4	1.0	5	4	US-09-187-859-187A	Sequence 187A, App1
64	4	1.0	5	4	US-09-187-859-188A	Sequence 188A, App1
65	4	1.0	5	4	US-09-187-859-189A	Sequence 189A, App1
66	4	1.0	5	4	US-09-187-859-190A	Sequence 190A, App1
67	4	1.0	5	4	US-09-187-859-191A	Sequence 191A, App1
68	4	1.0	5	4	US-09-187-859-192A	Sequence 192A, App1
69	4	1.0	5	4	US-09-187-859-193A	Sequence 193A, App1
70	4	1.0	5	4	US-09-187-859-194A	Sequence 194A, App1
71	4	1.0	5	4	US-09-187-859-195A	Sequence 195A, App1
72	4	1.0	5	4	US-09-187-859-196A	Sequence 196A, App1
73	4	1.0	5	4	US-09-187-859-197A	Sequence 197A, App1
74	4	1.0	5	4	US-09-187-859-198A	Sequence 198A, App1
75	4	1.0	5	4	US-09-187-859-199A	Sequence 199A, App1
76	4	1.0	5	4	US-09-187-859-200A	Sequence 200A, App1
77	4	1.0	5	4	US-09-187-859-201A	Sequence 201A, App1
78	4	1.0	5	4	US-09-187-859-202A	Sequence 202A, App1
79	4	1.0	5	4	US-09-187-859-203A	Sequence 203A, App1
80	4	1.0	5	4	US-09-187-859-204A	Sequence 204A, App1
81	4	1.0	5	4	US-09-187-859-205A	Sequence 205A, App1
82	4	1.0	5	4	US-09-187-859-206A	Sequence 206A, App1
83	4	1.0	5	4	US-09-187-859-207A	Sequence 207A, App1
84	4	1.0	5	4	US-09-187-859-208A	Sequence 208A, App1
85	4	1.0	5	4	US-09-187-859-209A	Sequence 209A, App1
86	4	1.0	5	4	US-09-187-859-210A	Sequence 210A, App1
87	4	1.0	5	4	US-09-187-859-211A	Sequence 211A, App1
88	4	1.0	5	4	US-09-187-859-212A	Sequence 212A, App1
89	4	1.0	5	4	US-09-187-859-213A	Sequence 213A, App1
90	4	1.0	5	4	US-09-187-859-214A	Sequence 214A, App1
91	4	1.0	5	4	US-09-187-859-215A	Sequence 215A, App1
92	4	1.0	5	4	US-09-187-859-216A	Sequence 216A, App1
93	4	1.0	5	4	US-09-187-859-217A	Sequence 217A, App1
94	4	1.0	5	4	US-09-187-859-218A	Sequence 218A, App1
95	4	1.0	5	4	US-09-187-859-219A	Sequence 219A, App1
96	4	1.0	5	4	US-09-187-859-220A	Sequence 220A, App1
97	4	1.0	5	4	US-09-187-859-221A	Sequence 221A, App1
98	4	1.0	5	4	US-09-187-859-222A	Sequence 222A, App1
99	4	1.0	5	4	US-09-187-859-223A	Sequence 223A, App1
100	4	1.0	5	4	US-09-187-859-224A	Sequence 224A, App1

[illegible]

247	3	0.7	4	1	US-08-076-002-41	Sequence 41, Appl	320	3	0.7	4	1	US-08-129-820-19	Sequence 27, Appl
248	3	0.7	4	1	US-08-243-873-1	Sequence 1, Appl	321	3	0.7	4	1	US-08-129-820-19	Sequence 19, Appl
249	3	0.7	4	1	US-08-276-919-6	Sequence 6, Appl	322	3	0.7	4	1	US-08-129-820-26	Sequence 26, Appl
250	3	0.7	4	1	US-08-201-081-1	Sequence 1, Appl	323	3	0.7	4	1	US-08-129-820-27	Sequence 27, Appl
251	3	0.7	4	1	US-08-201-081-8	Sequence 8, Appl	324	3	0.7	4	1	US-08-129-820-40	Sequence 42, Appl
252	3	0.7	4	1	US-08-249-371-5	Sequence 5, Appl	325	3	0.7	4	1	US-08-129-820-46	Sequence 48, Appl
253	3	0.7	4	1	US-08-151-568-20	Sequence 20, Appl	326	3	0.7	4	1	US-08-129-820-55	Sequence 57, Appl
254	3	0.7	4	1	US-08-040-119-5	Sequence 5, Appl	327	3	0.7	4	1	US-08-129-820-81	Sequence 83, Appl
255	3	0.7	4	1	US-07-890-422P-37	Sequence 37, Appl	328	3	0.7	4	1	US-08-129-820-87	Sequence 89, Appl
256	3	0.7	4	1	US-08-331-183-7	Sequence 7, Appl	329	3	0.7	4	1	US-08-129-820-87	Sequence 89, Appl
257	3	0.7	4	1	US-08-331-183-9	Sequence 9, Appl	330	3	0.7	4	1	US-08-129-820-90	Sequence 90, Appl
258	3	0.7	4	1	US-08-331-183-21	Sequence 21, Appl	331	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
259	3	0.7	4	1	US-08-331-183-22	Sequence 22, Appl	332	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
260	3	0.7	4	1	US-07-909-305-61	Sequence 61, Appl	333	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
261	3	0.7	4	1	US-07-969-305-62	Sequence 62, Appl	334	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
262	3	0.7	4	1	US-08-469-305-62	Sequence 62, Appl	335	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
263	3	0.7	4	1	US-08-469-305-62	Sequence 62, Appl	336	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
264	3	0.7	4	1	US-08-213-897A-16	Sequence 16, Appl	337	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
265	3	0.7	4	1	US-08-471-052A-152	Sequence 152, Appl	338	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
266	3	0.7	4	1	US-08-335-198-17	Sequence 17, Appl	339	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
267	3	0.7	4	1	US-08-335-198-25	Sequence 25, Appl	340	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
268	3	0.7	4	1	US-08-335-198-31	Sequence 31, Appl	341	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
269	3	0.7	4	1	US-08-335-198-31	Sequence 31, Appl	342	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
270	3	0.7	4	1	US-08-280-443-39	Sequence 39, Appl	343	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
271	3	0.7	4	1	US-08-544-147A-8	Sequence 8, Appl	344	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
272	3	0.7	4	1	US-08-458-998A-1	Sequence 1, Appl	345	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
273	3	0.7	4	1	US-08-458-998A-1	Sequence 1, Appl	346	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
274	3	0.7	4	1	US-08-458-998A-7	Sequence 7, Appl	347	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
275	3	0.7	4	1	US-08-458-998A-9	Sequence 9, Appl	348	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
276	3	0.7	4	1	US-08-458-998A-12	Sequence 12, Appl	349	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
277	3	0.7	4	1	US-08-458-998A-15	Sequence 15, Appl	350	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
278	3	0.7	4	1	US-08-458-998A-16	Sequence 16, Appl	351	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
279	3	0.7	4	1	US-08-358-160-168	Sequence 168, Appl	352	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
280	3	0.7	4	1	US-08-358-160-168	Sequence 168, Appl	353	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
281	3	0.7	4	1	US-08-157-150-226	Sequence 226, Appl	354	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
282	3	0.7	4	1	US-08-157-150-226	Sequence 226, Appl	355	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
283	3	0.7	4	1	US-08-157-150-226	Sequence 226, Appl	356	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
284	3	0.7	4	1	US-08-142-439A-8	Sequence 8, Appl	357	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
285	3	0.7	4	1	US-08-142-439A-9	Sequence 9, Appl	358	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
286	3	0.7	4	1	US-08-584-579-12	Sequence 12, Appl	359	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
287	3	0.7	4	1	US-08-082-8198-34	Sequence 34, Appl	360	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
288	3	0.7	4	1	US-08-457-459-39	Sequence 39, Appl	361	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
289	3	0.7	4	1	US-08-057-971-71	Sequence 71, Appl	362	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
290	3	0.7	4	1	US-08-240-711-2	Sequence 2, Appl	363	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
291	3	0.7	4	1	US-08-137-627-14	Sequence 14, Appl	364	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
292	3	0.7	4	1	US-08-449-202-4	Sequence 4, Appl	365	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
293	3	0.7	4	1	US-08-312-188-5	Sequence 5, Appl	366	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
294	3	0.7	4	1	US-08-451-566-20	Sequence 20, Appl	367	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
295	3	0.7	4	1	US-08-128-225-4	Sequence 4, Appl	368	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
296	3	0.7	4	1	US-08-451-566-20	Sequence 20, Appl	369	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
297	3	0.7	4	1	US-08-451-566-20	Sequence 20, Appl	370	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
298	3	0.7	4	1	US-08-451-566-20	Sequence 20, Appl	371	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
299	3	0.7	4	1	US-08-451-566-20	Sequence 20, Appl	372	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
300	3	0.7	4	1	US-08-549-008-7	Sequence 7, Appl	373	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
301	3	0.7	4	1	US-08-549-008-9	Sequence 9, Appl	374	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
302	3	0.7	4	1	US-08-549-008-22	Sequence 22, Appl	375	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
303	3	0.7	4	1	US-08-549-008-23	Sequence 23, Appl	376	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
304	3	0.7	4	1	US-08-549-008-29	Sequence 29, Appl	377	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
305	3	0.7	4	1	US-08-698-675-17	Sequence 17, Appl	378	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
306	3	0.7	4	1	US-08-675-140-4	Sequence 4, Appl	379	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
307	3	0.7	4	1	US-08-425-069-18	Sequence 18, Appl	380	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
308	3	0.7	4	1	US-08-486-013-1	Sequence 1, Appl	381	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
309	3	0.7	4	1	US-08-548-540-113	Sequence 113, Appl	382	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
310	3	0.7	4	1	US-08-726-160-23	Sequence 23, Appl	383	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
311	3	0.7	4	1	US-08-443-890-6	Sequence 6, Appl	384	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
312	3	0.7	4	1	US-08-406-192-25	Sequence 25, Appl	385	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
313	3	0.7	4	1	US-08-406-192-34	Sequence 34, Appl	386	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
314	3	0.7	4	1	US-08-406-192-44	Sequence 44, Appl	387	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
315	3	0.7	4	1	US-08-483-415-16	Sequence 16, Appl	388	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
316	3	0.7	4	1	US-08-460-448-64	Sequence 64, Appl	389	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
317	3	0.7	4	1	US-08-589-011-4	Sequence 4, Appl	390	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
318	3	0.7	4	1	US-08-173-510K-104	Sequence 104, Appl	391	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl
319	3	0.7	4	1	US-08-329-820-1	Sequence 1, Appl	392	3	0.7	4	1	US-08-129-820-192	Sequence 194, Appl

[illegible]

539	3	0.7	4	US-09-231-797-14	Sequence 14, Appl	612	3	0.7	4	US-09-056-226-16	Sequence 14, Appl
540	3	0.7	4	US-09-008-481A-1	Sequence 1, Appl	613	3	0.7	4	US-09-082-593-14	Sequence 14, Appl
541	3	0.7	4	US-09-198-209-3	Sequence 3, Appl	614	3	0.7	4	US-09-261-851-5	Sequence 7, Appl
542	3	0.7	4	US-08-851-812A-22	Sequence 22, Appl	615	3	0.7	4	US-09-261-851-7	Sequence 7, Appl
543	3	0.7	4	US-08-912-272-16	Sequence 16, Appl	616	3	0.7	4	US-08-515-681-3	Sequence 3, Appl
544	3	0.7	4	US-08-633-410-2	Sequence 2, Appl	617	3	0.7	4	US-09-142-699-1	Sequence 1, Appl
545	3	0.7	4	US-08-633-410-7	Sequence 7, Appl	618	3	0.7	4	US-09-142-699-2	Sequence 2, Appl
546	3	0.7	4	US-08-633-410-10	Sequence 10, Appl	619	3	0.7	4	US-08-478-029A-1	Sequence 1, Appl
547	3	0.7	4	US-08-633-410-15	Sequence 15, Appl	620	3	0.7	4	US-08-478-029A-4	Sequence 4, Appl
548	3	0.7	4	US-08-633-410-18	Sequence 18, Appl	621	3	0.7	4	US-09-058-562-6	Sequence 6, Appl
549	3	0.7	4	US-08-633-410-22	Sequence 22, Appl	622	3	0.7	4	US-08-974-462-17	Sequence 17, Appl
550	3	0.7	4	US-08-934-224-14	Sequence 14, Appl	623	3	0.7	4	US-08-974-462-18	Sequence 18, Appl
551	3	0.7	4	US-08-948-897-4	Sequence 4, Appl	624	3	0.7	4	US-08-974-462-19	Sequence 19, Appl
552	3	0.7	4	US-08-814-946-4	Sequence 4, Appl	625	3	0.7	4	US-08-974-462-20	Sequence 20, Appl
553	3	0.7	4	US-08-836-480-2	Sequence 2, Appl	626	3	0.7	4	US-09-181-919-91	Sequence 91, Appl
554	3	0.7	4	US-08-836-480-3	Sequence 3, Appl	627	3	0.7	4	US-08-182-919-91	Sequence 91, Appl
555	3	0.7	4	US-08-836-480-32	Sequence 32, Appl	628	3	0.7	4	US-09-224-681-84	Sequence 84, Appl
556	3	0.7	4	US-09-082-039A-3	Sequence 3, Appl	629	3	0.7	4	US-09-115-395-24	Sequence 24, Appl
557	3	0.7	4	US-08-477-781A-5	Sequence 5, Appl	630	3	0.7	4	US-08-436-729A-14	Sequence 14, Appl
558	3	0.7	4	US-09-222-373-52	Sequence 52, Appl	631	3	0.7	4	US-08-436-729A-161	Sequence 161, Appl
559	3	0.7	4	US-08-304-748-28	Sequence 28, Appl	632	3	0.7	4	US-08-515-681-5	Sequence 5, Appl
560	3	0.7	4	US-08-904-760A-30	Sequence 30, Appl	633	3	0.7	4	US-09-066-481-15	Sequence 15, Appl
561	3	0.7	4	US-08-933-843-14	Sequence 14, Appl	634	3	0.7	4	US-09-066-481-15	Sequence 15, Appl
562	3	0.7	4	US-08-950-860-14	Sequence 14, Appl	635	3	0.7	4	US-09-562-897-3	Sequence 3, Appl
563	3	0.7	4	US-08-950-860-15	Sequence 15, Appl	636	3	0.7	4	US-09-562-897-3	Sequence 3, Appl
564	3	0.7	4	US-08-581-662-9	Sequence 9, Appl	637	3	0.7	4	US-09-562-897-32	Sequence 32, Appl
565	3	0.7	4	US-08-581-662-12	Sequence 12, Appl	638	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
566	3	0.7	4	US-08-581-662-13	Sequence 13, Appl	639	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
567	3	0.7	4	US-08-581-662-23	Sequence 23, Appl	640	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
568	3	0.7	4	US-09-101-167-26	Sequence 26, Appl	641	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
569	3	0.7	4	US-09-101-167-35	Sequence 35, Appl	642	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
570	3	0.7	4	US-09-101-167-36	Sequence 36, Appl	643	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
571	3	0.7	4	US-09-101-167-38	Sequence 38, Appl	644	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
572	3	0.7	4	US-08-295-802-23	Sequence 23, Appl	645	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
573	3	0.7	4	US-08-963-1680-29	Sequence 29, Appl	646	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
574	3	0.7	4	US-08-822-324-22	Sequence 22, Appl	647	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
575	3	0.7	4	US-08-822-324-22	Sequence 22, Appl	648	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
576	3	0.7	4	US-09-039-657-24	Sequence 24, Appl	649	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
577	3	0.7	4	US-08-475-411A-1	Sequence 1, Appl	650	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
578	3	0.7	4	US-08-475-411A-2	Sequence 2, Appl	651	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
579	3	0.7	4	US-09-196-566A-17	Sequence 17, Appl	652	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
580	3	0.7	4	US-08-435-568A-17	Sequence 17, Appl	653	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
581	3	0.7	4	US-08-435-568A-18	Sequence 18, Appl	654	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
582	3	0.7	4	US-09-339-970-7	Sequence 7, Appl	655	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
583	3	0.7	4	US-09-339-970-16	Sequence 16, Appl	656	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
584	3	0.7	4	US-09-339-970-17	Sequence 17, Appl	657	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
585	3	0.7	4	US-09-339-970-18	Sequence 18, Appl	658	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
586	3	0.7	4	US-09-339-970-19	Sequence 19, Appl	659	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
587	3	0.7	4	US-09-339-970-34	Sequence 34, Appl	660	3	0.7	4	US-09-171-654-16	Sequence 16, Appl
588	3	0.7	4	US-08-934-223-14	Sequence 14, Appl	661	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
589	3	0.7	4	US-09-402-456-1	Sequence 1, Appl	662	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
590	3	0.7	4	US-08-859-242-11	Sequence 11, Appl	663	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
591	3	0.7	4	US-08-859-242-12	Sequence 12, Appl	664	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
592	3	0.7	4	US-08-859-242-22	Sequence 22, Appl	665	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
593	3	0.7	4	US-08-859-242-23	Sequence 23, Appl	666	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
594	3	0.7	4	US-08-859-242-29	Sequence 29, Appl	667	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
595	3	0.7	4	US-08-859-242-31	Sequence 31, Appl	668	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
596	3	0.7	4	US-08-859-242-32	Sequence 32, Appl	669	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
597	3	0.7	4	US-08-859-242-40	Sequence 40, Appl	670	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
598	3	0.7	4	US-08-859-242-42	Sequence 42, Appl	671	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
599	3	0.7	4	US-08-859-242-43	Sequence 43, Appl	672	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
600	3	0.7	4	US-08-859-242-58	Sequence 58, Appl	673	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
601	3	0.7	4	US-08-564-164A-15	Sequence 15, Appl	674	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
602	3	0.7	4	US-08-564-164A-17	Sequence 17, Appl	675	3	0.7	4	US-08-960-054A-21	Sequence 21, Appl
603	3	0.7	4	US-09-124-671-40	Sequence 40, Appl	676	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
604	3	0.7	4	US-08-974-549A-71	Sequence 71, Appl	677	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
605	3	0.7	4	US-08-974-549A-161	Sequence 161, Appl	678	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
606	3	0.7	4	US-08-974-549A-175	Sequence 175, Appl	679	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
607	3	0.7	4	US-08-974-549A-341	Sequence 341, Appl	680	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
608	3	0.7	4	US-09-103-663-17	Sequence 17, Appl	681	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
609	3	0.7	4	US-09-103-663-21	Sequence 21, Appl	682	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
610	3	0.7	4	US-09-142-078-33	Sequence 33, Appl	683	3	0.7	4	US-09-309-592-1	Sequence 1, Appl
611	3	0.7	4	US-08-924-002-19	Sequence 19, Appl	684	3	0.7	4	US-09-309-592-1	Sequence 1, Appl

831	3	0.7	5	1	US-07-657-7698-66	Sequence 66, App1	904	3	0.7	5	1	US-08-297-430-5	Sequence 5, App1	904	3	0.7	5	1	US-08-297-430-5	Sequence 5, App1
832	3	0.7	5	1	US-07-657-7698-67	Sequence 67, App1	905	3	0.7	5	1	US-08-297-430-6	Sequence 6, App1	905	3	0.7	5	1	US-08-297-430-6	Sequence 6, App1
833	3	0.7	5	1	US-07-801-388-4	Sequence 4, App11	906	3	0.7	5	1	US-08-297-430-15	Sequence 15, App1	906	3	0.7	5	1	US-08-297-430-15	Sequence 15, App1
834	3	0.7	5	1	US-07-924-753-6	Sequence 6, App11	907	3	0.7	5	1	US-08-297-430-24	Sequence 24, App1	907	3	0.7	5	1	US-08-297-430-24	Sequence 24, App1
835	3	0.7	5	1	US-07-989-962-6	Sequence 6, App11	908	3	0.7	5	1	US-08-297-430-25	Sequence 25, App1	908	3	0.7	5	1	US-08-297-430-25	Sequence 25, App1
836	3	0.7	5	1	US-07-989-962-7	Sequence 7, App11	909	3	0.7	5	1	US-08-297-430-26	Sequence 26, App1	909	3	0.7	5	1	US-08-297-430-26	Sequence 26, App1
837	3	0.7	5	1	US-07-838-410-9	Sequence 9, App11	910	3	0.7	5	1	US-08-297-430-27	Sequence 27, App1	910	3	0.7	5	1	US-08-297-430-27	Sequence 27, App1
838	3	0.7	5	1	US-08-061-065-4	Sequence 4, App11	911	3	0.7	5	1	US-08-297-430-28	Sequence 28, App1	911	3	0.7	5	1	US-08-297-430-28	Sequence 28, App1
839	3	0.7	5	1	US-07-728-2218-9	Sequence 8, App11	912	3	0.7	5	1	US-08-297-430-29	Sequence 29, App1	912	3	0.7	5	1	US-08-297-430-29	Sequence 29, App1
840	3	0.7	5	1	US-07-989-920-1	Sequence 1, App11	913	3	0.7	5	1	US-08-297-430-30	Sequence 30, App1	913	3	0.7	5	1	US-08-297-430-30	Sequence 30, App1
841	3	0.7	5	1	US-07-973-852-5	Sequence 5, App11	914	3	0.7	5	1	US-08-297-430-31	Sequence 31, App1	914	3	0.7	5	1	US-08-297-430-31	Sequence 31, App1
842	3	0.7	5	1	US-08-029-333-25	Sequence 25, App1	915	3	0.7	5	1	US-08-297-430-32	Sequence 32, App1	915	3	0.7	5	1	US-08-297-430-32	Sequence 32, App1
843	3	0.7	5	1	US-08-093-781-8	Sequence 8, App11	916	3	0.7	5	1	US-08-297-430-33	Sequence 33, App1	916	3	0.7	5	1	US-08-297-430-33	Sequence 33, App1
844	3	0.7	5	1	US-08-093-858-6	Sequence 6, App11	917	3	0.7	5	1	US-08-297-430-34	Sequence 34, App1	917	3	0.7	5	1	US-08-297-430-34	Sequence 34, App1
845	3	0.7	5	1	US-07-729-099-8	Sequence 8, App11	918	3	0.7	5	1	US-08-297-430-35	Sequence 35, App1	918	3	0.7	5	1	US-08-297-430-35	Sequence 35, App1
846	3	0.7	5	1	US-07-729-099-9	Sequence 9, App11	919	3	0.7	5	1	US-08-297-430-36	Sequence 36, App1	919	3	0.7	5	1	US-08-297-430-36	Sequence 36, App1
847	3	0.7	5	1	US-07-729-099-10	Sequence 10, App11	920	3	0.7	5	1	US-08-297-430-37	Sequence 37, App1	920	3	0.7	5	1	US-08-297-430-37	Sequence 37, App1
848	3	0.7	5	1	US-07-741-2130-20	Sequence 20, App1	921	3	0.7	5	1	US-08-297-430-38	Sequence 38, App1	921	3	0.7	5	1	US-08-297-430-38	Sequence 38, App1
849	3	0.7	5	1	US-07-741-2130-36	Sequence 36, App1	922	3	0.7	5	1	US-08-297-430-39	Sequence 39, App1	922	3	0.7	5	1	US-08-297-430-39	Sequence 39, App1
850	3	0.7	5	1	US-07-950-773-3	Sequence 3, App11	923	3	0.7	5	1	US-08-297-430-40	Sequence 40, App1	923	3	0.7	5	1	US-08-297-430-40	Sequence 40, App1
851	3	0.7	5	1	US-07-880-216-1	Sequence 1, App11	924	3	0.7	5	1	US-08-297-430-41	Sequence 41, App1	924	3	0.7	5	1	US-08-297-430-41	Sequence 41, App1
852	3	0.7	5	1	US-07-841-997A-11	Sequence 11, App1	925	3	0.7	5	1	US-08-297-430-42	Sequence 42, App1	925	3	0.7	5	1	US-08-297-430-42	Sequence 42, App1
853	3	0.7	5	1	US-07-805-727-5	Sequence 5, App11	926	3	0.7	5	1	US-08-297-430-43	Sequence 43, App1	926	3	0.7	5	1	US-08-297-430-43	Sequence 43, App1
854	3	0.7	5	1	US-07-805-727-10	Sequence 10, App1	927	3	0.7	5	1	US-08-297-430-44	Sequence 44, App1	927	3	0.7	5	1	US-08-297-430-44	Sequence 44, App1
855	3	0.7	5	1	US-07-805-727-20	Sequence 20, App1	928	3	0.7	5	1	US-08-297-430-45	Sequence 45, App1	928	3	0.7	5	1	US-08-297-430-45	Sequence 45, App1
856	3	0.7	5	1	US-08-055-988-6	Sequence 6, App11	929	3	0.7	5	1	US-08-297-430-46	Sequence 46, App1	929	3	0.7	5	1	US-08-297-430-46	Sequence 46, App1
857	3	0.7	5	1	US-08-106-509-5	Sequence 5, App11	930	3	0.7	5	1	US-08-297-430-47	Sequence 47, App1	930	3	0.7	5	1	US-08-297-430-47	Sequence 47, App1
858	3	0.7	5	1	US-08-271-698-1	Sequence 1, App11	931	3	0.7	5	1	US-08-297-430-48	Sequence 48, App1	931	3	0.7	5	1	US-08-297-430-48	Sequence 48, App1
859	3	0.7	5	1	US-07-665-927-7	Sequence 7, App11	932	3	0.7	5	1	US-08-297-430-49	Sequence 49, App1	932	3	0.7	5	1	US-08-297-430-49	Sequence 49, App1
860	3	0.7	5	1	US-09-384-618-7	Sequence 7, App11	933	3	0.7	5	1	US-08-297-430-50	Sequence 50, App1	933	3	0.7	5	1	US-08-297-430-50	Sequence 50, App1
861	3	0.7	5	1	US-09-384-618-7	Sequence 7, App11	934	3	0.7	5	1	US-08-297-430-51	Sequence 51, App1	934	3	0.7	5	1	US-08-297-430-51	Sequence 51, App1
862	3	0.7	5	1	US-09-384-618-7	Sequence 7, App11	935	3	0.7	5	1	US-08-297-430-52	Sequence 52, App1	935	3	0.7	5	1	US-08-297-430-52	Sequence 52, App1
863	3	0.7	5	1	US-07-969-307A-7	Sequence 7, App11	936	3	0.7	5	1	US-08-297-430-53	Sequence 53, App1	936	3	0.7	5	1	US-08-297-430-53	Sequence 53, App1
864	3	0.7	5	1	US-07-729-099-8	Sequence 8, App11	937	3	0.7	5	1	US-08-297-430-54	Sequence 54, App1	937	3	0.7	5	1	US-08-297-430-54	Sequence 54, App1
865	3	0.7	5	1	US-07-729-099-9	Sequence 9, App11	938	3	0.7	5	1	US-08-297-430-55	Sequence 55, App1	938	3	0.7	5	1	US-08-297-430-55	Sequence 55, App1
866	3	0.7	5	1	US-09-257-782-3	Sequence 3, App11	939	3	0.7	5	1	US-08-297-430-56	Sequence 56, App1	939	3	0.7	5	1	US-08-297-430-56	Sequence 56, App1
867	3	0.7	5	1	US-07-671-757-4	Sequence 4, App11	940	3	0.7	5	1	US-08-297-430-57	Sequence 57, App1	940	3	0.7	5	1	US-08-297-430-57	Sequence 57, App1
868	3	0.7	5	1	US-07-671-757-15	Sequence 15, App1	941	3	0.7	5	1	US-08-297-430-58	Sequence 58, App1	941	3	0.7	5	1	US-08-297-430-58	Sequence 58, App1
869	3	0.7	5	1	US-07-671-757-19	Sequence 19, App1	942	3	0.7	5	1	US-08-297-430-59	Sequence 59, App1	942	3	0.7	5	1	US-08-297-430-59	Sequence 59, App1
870	3	0.7	5	1	US-07-671-757-25	Sequence 25, App1	943	3	0.7	5	1	US-08-297-430-60	Sequence 60, App1	943	3	0.7	5	1	US-08-297-430-60	Sequence 60, App1
871	3	0.7	5	1	US-07-671-757-35	Sequence 35, App1	944	3	0.7	5	1	US-08-297-430-61	Sequence 61, App1	944	3	0.7	5	1	US-08-297-430-61	Sequence 61, App1
872	3	0.7	5	1	US-08-300-372-5	Sequence 5, App11	945	3	0.7	5	1	US-08-297-430-62	Sequence 62, App1	945	3	0.7	5	1	US-08-297-430-62	Sequence 62, App1
873	3	0.7	5	1	US-08-300-372-10	Sequence 10, App1	946	3	0.7	5	1	US-08-297-430-63	Sequence 63, App1	946	3	0.7	5	1	US-08-297-430-63	Sequence 63, App1
874	3	0.7	5	1	US-08-300-372-20	Sequence 20, App1	947	3	0.7	5	1	US-08-297-430-64	Sequence 64, App1	947	3	0.7	5	1	US-08-297-430-64	Sequence 64, App1
875	3	0.7	5	1	US-08-234-602-12	Sequence 12, App1	948	3	0.7	5	1	US-08-297-430-65	Sequence 65, App1	948	3	0.7	5	1	US-08-297-430-65	Sequence 65, App1
876	3	0.7	5	1	US-08-432-617-5	Sequence 5, App11	949	3	0.7	5	1	US-08-297-430-66	Sequence 66, App1	949	3	0.7	5	1	US-08-297-430-66	Sequence 66, App1
877	3	0.7	5	1	US-08-081-539-107	Sequence 107, App1	950	3	0.7	5	1	US-08-297-430-67	Sequence 67, App1	950	3	0.7	5	1	US-08-297-430-67	Sequence 67, App1
878	3	0.7	5	1	US-08-139-089-2	Sequence 2, App11	951	3	0.7	5	1	US-08-297-430-68	Sequence 68, App1	951	3	0.7	5	1	US-08-297-430-68	Sequence 68, App1
879	3	0.7	5	1	US-08-129-089-3	Sequence 3, App11	952	3	0.7	5	1	US-08-297-430-69	Sequence 69, App1	952	3	0.7	5	1	US-08-297-430-69	Sequence 69, App1
880	3	0.7	5	1	US-08-129-089-4	Sequence 4, App11	953	3	0.7	5	1	US-08-297-430-70	Sequence 70, App1	953	3	0.7	5	1	US-08-297-430-70	Sequence 70, App1
881	3	0.7	5	1	US-08-129-089-14	Sequence 14, App1	954	3	0.7	5	1	US-08-297-430-71	Sequence 71, App1	954	3	0.7	5	1	US-08-297-430-71	Sequence 71, App1
882	3	0.7	5	1	US-08-129-089-15	Sequence 15, App1	955	3	0.7	5	1	US-08-297-430-72	Sequence 72, App1	955	3	0.7	5	1	US-08-297-430-72	Sequence 72, App1
883	3	0.7	5	1	US-08-129-089-20	Sequence 20, App1	956	3	0.7	5	1	US-08-297-430-73	Sequence 73, App1	956	3	0.7	5	1	US-08-297-430-73	Sequence 73, App1
884	3	0.7	5	1	US-08-129-089-21	Sequence 21, App1	957	3	0.7	5	1	US-08-297-430-74	Sequence 74, App1	957	3	0.7	5	1	US-08-297-430-74	Sequence 74, App1
885	3	0.7	5	1	US-08-129-089-22	Sequence 22, App1	958	3	0.7	5	1	US-08-297-430-75	Sequence 75, App1	958	3	0.7	5	1	US-08-297-430-75	Sequence 75, App1
886	3	0.7	5	1	US-08-014-979-36	Sequence 36, App1	959	3	0.7	5	1	US-08-297-430-76	Sequence 76, App1	959	3	0.7	5	1	US-08-297-430-76	Sequence 76, App1
887	3	0.7	5	1	US-07-609-716-2	Sequence 2, App11	960	3	0.7	5	1	US-08-297-430-77	Sequence 77, App1	960	3	0.7	5	1	US-08-297-430-77	Sequence 77, App1
888	3	0.7	5	1	US-08-083-948-12	Sequence 12, App1	961	3	0.7	5	1	US-08-297-430-78	Sequence 78, App1	961	3	0.7	5	1	US-08-297-430-78	Sequence 78, App1
889	3	0.7	5	1	US-08-199-778-1	Sequence 1, App11	962	3	0.7	5	1	US-08-297-430-79	Sequence 79, App1	962	3	0.7	5	1	US-08-297-430-79	Sequence 79, App1
890	3	0.7	5	1	US-07-972-087-20	Sequence 20, App1	963	3	0.7	5	1	US-08-297-430-80	Sequence 80, App1	963	3	0.7	5	1	US-08-297-430-80	Sequence 80, App1
891	3	0.7	5	1	US-08-031-603A-20	Sequence 20, App1	964	3	0.7	5	1	US-08-297-430-81	Sequence 81, App1	964	3	0.7	5	1	US-08-297-430-81	Sequence 81, App1
892	3	0.7	5	1	US-08-385-443-6	Sequence 6, App11	965	3	0.7	5										


```

GENERAL INFORMATION:
APPLICANT: STEPHENS, GERO J.
APPLICANT: WENDEL, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL, WILHELM, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIOFUNCTIONAL PROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROUGH
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESS: Evenson, McGowan, Edwards & Lavanhan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09/499,741
FILING DATE: 20-JUL-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41445
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-043-741-72

Query Match 1.24; Score 5; DB 1; Length 6;
Best Local Similarity 100.0%; Prod. No. 1.70e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 PKAY 290
DB 2 PKAY 6

```

```

APPLICATION NUMBER: US-08/486,721A
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 06/091,991
FILING DATE: 12-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: BATH, Richard S.
REGISTRATION NUMBER: 28,180
REFERENCE/DOCKET NUMBER: 148/41445
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 319-4400
TELEFAX: (212) 319-5101
INDEX: 26268
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
US-08-486-721A-19

Query Match 1.24; Score 5; DB 1; Length 6;
Best Local Similarity 100.0%; Prod. No. 1.70e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NWEVE 9
DB 1 NWEVE 9

RESULT 5
US-08-720-012-72
Sequence 72; Application 05/08720012
Patent No. 5747291
GENERAL INFORMATION:
APPLICANT: STEPHENS, GERO J.
APPLICANT: WENDEL, STEPHAN
APPLICANT: SCHNEIDER, JOHANNES
APPLICANT: HEINZEL, WILHELM, REGINA
APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIOFUNCTIONAL PROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROUGH
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESS: Evenson, McGowan, Edwards & Lavanhan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 06/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41445
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 72:

```


STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,833B
FILING DATE: 29-MAR-1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 076863/1995
FILING DATE: 31-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 271196/1995
FILING DATE: 19-OCT-1995
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-623-833B-67

Query Match 1.00; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Prod. No. 1.7e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 155 DDER 158
DB 1 DDER 4

RESULT 9
US-08-392-338A-7
Sequence 7, Application US/08392338A
Patent No. 5849620
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James P.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filippa, David
TITLE OF INVENTION: Multivalent Antigen-Binding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Stereo, Kossler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,338A
FILING DATE: 22-FEB-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/989,846
FILING DATE: 20-NOV-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/796,936
FILING DATE: 25-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
PREFPREF/DOCPFT NUMBER: 0077 0010007

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 471-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-392-338A-7

Query Match 1.00; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Prod. No. 1.7e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 52 VTWS 55
DB 1 VTWS 4

RESULT 10
US-08-242-549D-52
Sequence 52, Application US/08242549D
Patent No. 5965709
GENERAL INFORMATION:
APPLICANT: Presto, Leonard G.
APPLICANT: Jaretto, Paula M.
TITLE OF INVENTION: Ige Antagonists
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinWin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,539D
FILING DATE: 21-APR-1994
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/178584
FILING DATE: 07-JAN-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/744768
FILING DATE: 14-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
TELEPHONE/TELEFAX NUMBER: 1071894
TELEPHONE: 650/225-1489
TELEFAX: 650/225-1489
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-242-549D-52

Query Match 1.00; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Prod. No. 1.7e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 207 MALE 210
DB 1 MALE 4

SOFTWARE Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/172,014
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/292,228
 FILING DATE: 22-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/989,846
 FILING DATE: 20-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/796,936
 FILING DATE: 25-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldstein, Jorge A.
 REGISTRATION NUMBER: 29,021
 REFERENCE/AGENT NUMBER: 0977,003000P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2640
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 US-09-172-014-7

Query Match 1.08; Score 4; DB 3; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 52 VTVS 55
 1111
 DB 1 VTVS 4

RESULT 14

US-09-166-094-7
 Sequence 7, Application US/09166094
 Patent No. 6121424
 GENERAL INFORMATION:
 APPLICANT: Whitlow, Marc
 APPLICANT: Wood, James F.
 APPLICANT: Hartman, Karl
 APPLICANT: Bird, Robert
 APPLICANT: Filpula, David
 APPLICANT: Kollence, Michelle
 TITLE OF INVENTION: Multivalent Antigen Binding Proteins
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.,
 STREET: 1100 New York Avenue, NW
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS-MS DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/166,094
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/392,338
 FILING DATE: 22-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/989,846
 FILING DATE: 20-NOV-1992
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/796,936
 FILING DATE: 25-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldstein, Jorge A.
 REGISTRATION NUMBER: 29,021
 REFERENCE/AGENT NUMBER: 0977,003000A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2640
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 US-09-166-094-7

Query Match 1.08; Score 4; DB 3; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0;

QY 52 VTVS 55
 1111
 DB 1 VTVS 4

RESULT 15
 US-09-236-160-54
 Sequence 54, Application US/09236160A

Patent No. 6307090
 GENERAL INFORMATION:
 APPLICANT: Burke Jr., Terence R.
 TITLE OF INVENTION: ANTAGONIST ELIMINATING BIRIVALLIES HAVING THE SIGNAL
 FILE REFERENCE: 175895
 CURRENT APPLICATION NUMBER: 09/236,160A
 CURRENT FILING DATE: 1999-01-22
 NUMBER OF SEQ ID NOS: 55
 SOFTWARE: Patent in Ver. 2.1
 SEQ ID NO 54
 LENGTH: 4
 TYPE: PRT
 ORGANISM: mammalian
 US-09-236-160-54

Query Match 1.08; Score 4; DB 4; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 INOS 400
 1111
 DB 1 INOS 4

RESULT 16

US-09-5523089-9
 Patent No. 5523089
 APPLICANT: BERKSTEIN, AVN, HARRIS, ALAN G., MANTAKIS, DIMITRIOS A.
 TITLE OF INVENTION: BOREHOLA ANTICEN
 NUMBER OF SEQUENCES: 38
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/552,309
 FILING DATE: 22-JUN-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 924,798
 FILING DATE: 06-AUG-1992
 APPLICATION NUMBER: 422,881
 FILING DATE: 18-OCT-1989
 SEQ ID NO:9:
 LENGTH: 4
 US-09-5523089-9

Query Match 1.08; Score 4; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Message: 1, Subject: [REDACTED], Message-ID: [REDACTED], Date: [REDACTED]

27 42 1000
14 1 1000 1

RESULT 17

09:06:456.040.04

2 Sequence 44, Application: us/09040400

2 Patient No: 0904000

2 GENERAL INFORMATION

2 APPLICANT: [REDACTED]

2 APPLICANT: [REDACTED]

2 TITLE OF INVENTION: [REDACTED]

2 NUMBER OF SEQUENCES: 4

2 ADDRESS: [REDACTED]

2 STREET: [REDACTED]

2 CITY: [REDACTED]

2 STATE: [REDACTED]

2 COUNTRY: US

2 OTHER REFERENCE NO.

2 METHOD TYPE: [REDACTED]

2 SEQUENCE: 1001, Name of the

2 OPERATING SYSTEM: [REDACTED]

2 SOFTWARE: [REDACTED]

2 CURRENT APPLICATION DATA

2 APPLICATION NUMBER: [REDACTED]

2 FILING DATE: [REDACTED]

2 CLASSIFICATION: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 GENERAL INFORMATION

2 APPLICANT: [REDACTED]

2 APPLICANT: [REDACTED]

2 TITLE OF INVENTION: [REDACTED]

2 NUMBER OF SEQUENCES: 4

2 ADDRESS: [REDACTED]

2 STREET: [REDACTED]

2 CITY: [REDACTED]

2 STATE: [REDACTED]

2 COUNTRY: US

2 OTHER REFERENCE NO.

2 METHOD TYPE: [REDACTED]

2 SEQUENCE: 1001, Name of the

2 OPERATING SYSTEM: [REDACTED]

2 SOFTWARE: [REDACTED]

2 CURRENT APPLICATION DATA

2 APPLICATION NUMBER: [REDACTED]

2 FILING DATE: [REDACTED]

2 CLASSIFICATION: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

2 APPLICANT: [REDACTED]

2 FILING DATE: [REDACTED]

CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: MISTOCK, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7005-077
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-211-747-6

Query Match: 1.0%; Score 4; DB 1; Length 5;
 Best Local Similarity: 100.0%; Prod. No. 1.7e-05;
 Matches: 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 205 AVAA 208
 DB 1 AVAA 4

RESULT 20
 US-08-486-721A-20
 Sequence 20, Application US/08/486-721A
 Patent No. 5734025
 GENERAL INFORMATION:
 APPLICANT: FUKASAWA, Chikafusa
 TITLE OF INVENTION: Method of Producing
 TITLE OF INVENTION: Asparaginyl Endoprotease
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESS: Fishbait, Inc., Goodman, Langer & Chick
 STREET: 767 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10017-2023
 TELEPHONE: (212) 319-4900
 TELEFAX: (212) 319-5101
 TELEX: 246268
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-486-721A-20

Query Match: 1.0%; Score 4; DB 1; Length 5;
 Best Local Similarity: 100.0%; Prod. No. 1.7e-05;

Matches: 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 NVEE 9
 DB 1 NVEE 4

RESULT 21
 US-08-220-401-19
 Sequence 19, Application US/98220401
 Patent No. 5780266
 GENERAL INFORMATION:
 APPLICANT: Daley, Beverly
 APPLICANT: Yamahaka, Miles
 APPLICANT: Arce, William M.
 APPLICANT: Chavez Jr., Lloyd G.
 TITLE OF INVENTION: FELINE INFECTIOUS PERITONITIS VIRUS
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESS: MERRISON & FORKSTER
 STREET: 2000 Potomac/Valley Avenue, NW, Suite 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006-1012
 TELEPHONE: (202) 887-0763
 TELEFAX: (202) 887-1500
 TELEX: 90-4040 MRSNORSWSH
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-220-401-19

Query Match: 1.0%; Score 4; DB 1; Length 5;
 Best Local Similarity: 100.0%; Prod. No. 1.7e-05;
 Matches: 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 48 IVLV 51
 DB 2 IVLV 5

RESULT 22
 US-08-467-607-10
 Sequence 10, Application US/98467607
 Patent No. 578434
 GENERAL INFORMATION:
 APPLICANT: TUNG, JAY S.
 APPLICANT: SUNDA, SUKANTO
 APPLICANT: MCDONAGHUE, LISA
 APPLICANT: VANSON, GWEN
 APPLICANT: ANDERSON, JOHN

NAME/KEY: peptide
LOCATION: 1.5
OTHER INFORMATION: /note="Mch3"

US-08-665-220-25
Query Match 1.0%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.70-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 315 10AD 318
DB 1 10AD 4

RESULT 25
US-08-437-362-19
Sequence 19, Application US/08437362
Patent No. 581104
Patent No. 581104 5776469

GENERAL INFORMATION:
APPLICANT: Dale, Beverly
APPLICANT: Yamanaka, Miles
APPLICANT: Acree, William M.
APPLICANT: Chavez Jr., Lloyd G.
TITLE OF INVENTION: PEPTIDE EFFECTIVE PEPTIDE-TITLE VITEC
TITLE OF INVENTION: DIAGNOSTIC TOOLS

NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: MOPPSON & POKSTER
STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20006 1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,362
FILING DATE: 09-MAY-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,659
REFERENCE/DOCKET NUMBER: 13251-20001.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)887-1500
TELEFAX: (202)887-0763
TELEX: 90-4030 MRSNFOERSWSH
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-437-362-19

Query Match 1.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.70-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 1VLV 51
DB 2 1VLV 5

RESULT 26
US-08-469-362-10
Sequence 10, Application US/08469362
Patent No. 5849711
GENERAL INFORMATION:

APPLICANT: TONG, JAY S.
APPLICANT: STINA, SORANTO
APPLICANT: MCINTOSH, LISA
APPLICANT: SENG, CHRISTOPHER M.F.
TITLE OF INVENTION: NEW METHOD AND
TITLE OF INVENTION: COMPOSITIONS FOR TREATING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: ATHENA BIOSCIENCES
STREET: 800 F. Gateway Blvd.
CITY: South San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,362
FILING DATE: 06-JUN 1995
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: COVALL, JOSH M.
REGISTRATION NUMBER: 42,741
REFERENCE/DOCKET NUMBER: 002010-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE: Modified site
LOCATION: 5
OTHER INFORMATION: /note "This position is 1yr or
OTHER INFORMATION: Trp."

US-08-469-362-10

Query Match 1.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.70-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 CASC 200
DB 1 CASC 4

RESULT 27
US-08-618-408B-25
Sequence 25, Application US/08618408B
Patent No. 5851815
GENERAL INFORMATION:

APPLICANT: Almonte, Fred S.
APPLICANT: Fernandez-Almonte, Teresa
APPLICANT: Litwack, Gerald
APPLICANT: Armstrong, Robert
APPLICANT: Tomasselli, Kevin
TITLE OF INVENTION: Method and Means for Selective Absorption
TITLE OF INVENTION: Processes, Methods and Means for Selective Absorption and Methods of Use
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESS: Campbell and Piores
STREET: 4470 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:

Db 1 KSSR 4

RESULT 30

US-08-637-7598-81

Sequence 81: Application US/086177598

Patent No. 5876931

GENERAL INFORMATION:

APPLICANT: David William Holston

TITLE OF INVENTION: Identification of Genes

NUMBER OF SEQUENCES: 501

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637,7598

FILING DATE: 03-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02875

FILING DATE: 11-DEC-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patricia L.

REGISTRATION NUMBER: 31,284

REFERENCE/INVENT NUMBER: 101

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794

TELEFAX: (404) 873-8795

INFORMATION FOR SEQ ID NO: 81:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-08-637-7598-81

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 31

US-08-928-958-20

Sequence 20: Application US/08928958

Patent No. 5877282

GENERAL INFORMATION:

APPLICANT: NADLER, STEVEN G.

APPLICANT: CLEVELAND, JEFFREY S.

APPLICANT: BLAKE, JAMES

APPLICANT: HAPPA, OMAR F.

TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 81:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 32

US-08-742-440A-14

Sequence 14: Application US/0874440A

Patent No. 5892014

GENERAL INFORMATION:

APPLICANT: Conklin, Shaun

APPLICANT: Ishihara, Hiroaki

APPLICANT: Connelly, Andrew

TITLE OF INVENTION: Protease Activated Receptor

TITLE OF INVENTION: 3 and Uses thereof

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 33

US-08-928-958-20

Sequence 20: Application US/08928958

Patent No. 5877282

GENERAL INFORMATION:

APPLICANT: NADLER, STEVEN G.

APPLICANT: CLEVELAND, JEFFREY S.

APPLICANT: BLAKE, JAMES

APPLICANT: HAPPA, OMAR F.

TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 34

US-08-742-440A-14

Sequence 14: Application US/0874440A

Patent No. 5892014

GENERAL INFORMATION:

APPLICANT: Conklin, Shaun

APPLICANT: Ishihara, Hiroaki

APPLICANT: Connelly, Andrew

TITLE OF INVENTION: Protease Activated Receptor

TITLE OF INVENTION: 3 and Uses thereof

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 35

US-08-928-958-20

Sequence 20: Application US/08928958

Patent No. 5877282

GENERAL INFORMATION:

APPLICANT: NADLER, STEVEN G.

APPLICANT: CLEVELAND, JEFFREY S.

APPLICANT: BLAKE, JAMES

APPLICANT: HAPPA, OMAR F.

TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 36

US-08-742-440A-14

Sequence 14: Application US/0874440A

Patent No. 5892014

GENERAL INFORMATION:

APPLICANT: Conklin, Shaun

APPLICANT: Ishihara, Hiroaki

APPLICANT: Connelly, Andrew

TITLE OF INVENTION: Protease Activated Receptor

TITLE OF INVENTION: 3 and Uses thereof

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 37

US-08-928-958-20

Sequence 20: Application US/08928958

Patent No. 5877282

GENERAL INFORMATION:

APPLICANT: NADLER, STEVEN G.

APPLICANT: CLEVELAND, JEFFREY S.

APPLICANT: BLAKE, JAMES

APPLICANT: HAPPA, OMAR F.

TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 38

US-08-742-440A-14

Sequence 14: Application US/0874440A

Patent No. 5892014

GENERAL INFORMATION:

APPLICANT: Conklin, Shaun

APPLICANT: Ishihara, Hiroaki

APPLICANT: Connelly, Andrew

TITLE OF INVENTION: Protease Activated Receptor

TITLE OF INVENTION: 3 and Uses thereof

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KOHNS & ASSOCIATES

REGISTRATION NUMBER:

REFERENCE/INVENT NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE:

TELEFAX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-928-958-20

Query Match

Best Local Similarity: 100.0%; Pred. No. 1.7e+05;

Matches: 4; Conservative: 0; Mismatches: 0; Indels: 0;

Db 2 LAFU 5

248 LAFU 251

1111

RESULT 39

US-08-928-958-20

Sequence 20: Application US/08928958

Patent No. 5877282

GENERAL INFORMATION:

APPLICANT: NADLER, STEVEN G.

APPLICANT: CLEVELAND, JEFFREY S.

APPLICANT: BLAKE, JAMES

APPLICANT: HAPPA, OMAR F.

TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-755-728-6

Query Match 1.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 1SCP 17
1111
DB 2 1SCP 5

RESULT 35
US-09-072-429-20
Sequence 20, Application US/09072429
Patent No. 5962416
GENERAL INFORMATION:

APPLICANT: Nadler, Steven G.
TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
TITLE OF INVENTION: INHIBITOR OF REVERSE TRANSCRIPTASE AND AN
TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:

ADDRESSEE: Bristol-Myers Squibb Company
STREET: P.O. Box 4000
CITY: Princeton

STATE: New Jersey

COUNTRY: USA

ZIP: 08543-4000

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.40

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/072,429

FILING DATE: 04-MAY-1998

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Klein, Christopher A.

REGISTRATION NUMBER: 34,463

REFERENCE/DOCKET NUMBER: ON0141B

TELEPHONE: (609) 252-3714

TELEFAX: (609) 252-4526

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-072-429-20

Query Match 1.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 KKKY 93
1111
DB 1 KKKY 4

RESULT 36
US-08-974-655-6
Sequence 6, Application US/08974655
Patent No. 5972676

GENERAL INFORMATION:

APPLICANT: Plowman, Gregory

ADDRESSEE: Mossie, Kevin

TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AHR-1

TITLE OF INVENTION: AHR-1 AND 2 MUTATIONS AND METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 644 West Fifth Street

CITY: San Francisco

STATE: California

COUNTRY: U.S.A.

ZIP: 94071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 Mb

COMPUTER: IBM compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,655

FILING DATE:

CLASSIFICATION: 445

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Marburg, Richard J.

REGISTRATION NUMBER: 32,327

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955 0440

INFORMATION FOR SEQ ID NO: 63

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Peptide

US-08-974-655-6

Query Match 1.0%; Score 4; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.7e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 1SCP 17
1111
DB 2 1SCP 5

RESULT 37
US-08-814-567A-6
Sequence 6, Application US/08814567A
Patent No. 5988598

GENERAL INFORMATION:

APPLICANT: ANGLADE, ROY

ADDRESSEE: SULLIVAN, DANIEL M.

TITLE OF INVENTION: LAROCHELLE WILLIAM

TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF PREPARATION

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 FRANCHISE STREET, NE

CITY: ATLANTA

STATE: GEORGIA

COUNTRY: USA

ZIP: 30003-1911

COMPUTER READABLE FORM:


```

: GENERAL INFORMATION:
: APPLICANT: Green, Lawrence K.
: APPLICANT: Shackelford, Nicholas V.
: APPLICANT: Ivanov, Vadim T.
: APPLICANT: Mikhailova, Inessa I.
: APPLICANT: Vaskovsky, Boris V.
: APPLICANT: Mikhailov, Alexander N.
: APPLICANT: Khavinson, Vladimir K.
: APPLICANT: Morozov, Vyacheslav G.
: TITLE OF INVENTION: Immunomodulating peptides and methods
: TITLE OF INVENTION: of use
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 09/542,511
: FILING DATE: 28-OCT-1993
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/783,517
: FILING DATE: 28-OCT-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/816,205
: FILING DATE: 02-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/647,633
: FILING DATE: 28-OCT-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Storella, John R.
: REGISTRATION NUMBER: 32,944
: REFERENCE/WORKSHEET NUMBER: 1542-5-2
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ. ID NO.: 34:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-144-779C-34

```

```

Query Match 1.00; Score 4; DB 3; Length 5;
Best local similarity 100.0%; Prod. No. 175-05;
Matches 4; Conservative 0; Mismatches 0; Totals 0; Gaps 0;

```

```

QY 269 AEEK 272
DB 2 AEEK 5

```

```

: RESULT 41
: US-09-260-190-7
: Sequence 7: Application US/09260190
: Patent No. 6096713
: GENERAL INFORMATION:
: APPLICANT: Green, Lawrence K.
: APPLICANT: Blasecki, John W.
: TITLE OF INVENTION: Pharmaceutical Antiproliferative Peptide
: TITLE OF INVENTION: Compositions and Methods of Use thereof
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 09/542,511
: FILING DATE: 01-MAR-1993
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/548,701
: FILING DATE: 03-OCT-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/414,764
: FILING DATE: 13-MAR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Powers, Margaret A.
: REGISTRATION NUMBER: 49,804
: REFERENCE/WORKSHEET NUMBER: 01542-002-2-015
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0400
: INFORMATION FOR SEQ. ID NO.: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-260-190-7

```

```

Query Match 1.00; Score 4; DB 3; Length 5;
Best local similarity 100.0%; Prod. No. 175-05;
Matches 4; Conservative 0; Mismatches 0; Totals 0; Gaps 0;

```

```

QY 269 AEEK 272
DB 2 AEEK 5

```

```

: RESULT 42
: US-08-484-511-2
: Sequence 2: Application US/08484511
: Patent No. 6100480
: GENERAL INFORMATION:
: APPLICANT: Green, Lawrence K.
: APPLICANT: Shackelford, Nicholas V.
: APPLICANT: Ivanov, Vadim T.
: APPLICANT: Mikhailova, Inessa I.
: APPLICANT: Vaskovsky, Boris V.
: APPLICANT: Mikhailov, Alexander N.
: APPLICANT: Khavinson, Vladimir K.
: APPLICANT: Morozov, Vyacheslav G.
: TITLE OF INVENTION: Pharmaceutical lysine containing
: TITLE OF INVENTION: Polypeptide Compositions and Methods of Use thereof
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew
: STREET: Two Embarcadero Center, Eighth Floor, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```


SEQ ID NO 76
LENGTH: 5
TYPE: PRT
ORGANISM: Molluscum contagiosum virus
FEATURE:
OTHER INFORMATION: Moll IV of PNA quantity/translation
US-09-415-444-76

Query Match 1.0% Score 4; DB 4; Length 5;
Best local Similarity 100.0%; Prod. No. 1.7c-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 GVL 344
DB 2 GVL 5

RESULT 46
US-09-257-218-36
Sequence 36, Application US/09/257218
Patent No. 6271361
GENERAL INFORMATION:
APPLICANT: Almonti, Ema S.
APPLICANT: Fernandes-Almonti, Teresa
APPLICANT: Litwak, Gerald
TITLE OF INVENTION: Apoptotic Protease Mite, Nucleic Acids
TITLE OF INVENTION: Encoding Same and Methods of Use
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/257,218
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/865,579
FILING DATE: 29-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-ID 2180
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-9849
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-257-218-36

Query Match 1.0% Score 4; DB 4; Length 5;
Best local Similarity 100.0%; Prod. No. 1.7c-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 315 IQAD 318
DB 1 IQAD 4

RESULT 47
US-09-311-760-36

Sequence 36, Application US/09/411760
Patent No. 6274318
GENERAL INFORMATION:
APPLICANT: Almonti, Ema S.
APPLICANT: Fernandes-Almonti, Teresa
APPLICANT: Litwak, Gerald
TITLE OF INVENTION: Apoptotic Protease Mite, Nucleic Acids
TITLE OF INVENTION: Encoding Same and Methods of Use
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/311,760
FILING DATE: 13-May-1999
CLASSIFICATION: unknown
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/865,579
FILING DATE: unknown
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-ID 2180
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-9849
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-311-760-36

Query Match 1.0% Score 4; DB 4; Length 5;
Best local Similarity 100.0%; Prod. No. 1.7c-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 IQAD 318
DB 1 IQAD 4

RESULT 48
US-09-291-692-25
Sequence 25, Application US/09/291692
Patent No. 6287795
GENERAL INFORMATION:
APPLICANT: Almonti, Ema S.
APPLICANT: Fernandes-Almonti, Teresa
APPLICANT: Litwak, Gerald
APPLICANT: Armstrong, Robert
APPLICANT: Terasaki, Kevin
TITLE OF INVENTION: Moll IV of PNA quantity/translation
TITLE OF INVENTION: Nucleic Acids Encoding and Methods of Use
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104

[illegible][illegible]

523	4	1.0	8	14	AAK3580.0	Hepatitis C virus	596	4	1.0	8	17	AAK6194.9	Hepatitis C virus
524	4	1.0	8	14	AAK3584.4	Hepatitis C virus	597	4	1.0	8	16	AAK6171.3	Hepatitis C virus
525	4	1.0	8	14	AAK3581.2	Hepatitis C virus	598	4	1.0	8	18	AAK4849.9	Hepatitis C virus
526	4	1.0	8	14	AAK3584.4	Hepatitis C virus	599	4	1.0	8	18	AAK4742.6	Hepatitis C virus
527	4	1.0	8	14	AAK3582.7	Hepatitis C virus	600	4	1.0	8	18	AAK4825.3	Hepatitis C virus
528	4	1.0	8	14	AAK3582.4	Hepatitis C virus	601	4	1.0	8	18	AAK4109.7	Hepatitis C virus
529	4	1.0	8	14	AAK3587.8	Hepatitis C virus	602	4	1.0	8	18	AAK2657.5	Hepatitis C virus
530	4	1.0	8	14	AAK3366.9	Amphiphilic peptide	603	4	1.0	8	16	AAK2419.5	Hepatitis C virus
531	4	1.0	8	14	AAK3335.4	Sequence of tryptid	604	4	1.0	8	16	AAK2419.5	Hepatitis C virus
532	4	1.0	8	14	AAK3115.9	C-terminal subunit	605	4	1.0	8	18	AAK2420.1	Hepatitis C virus
533	4	1.0	8	14	AAK3537.7	Amphiphilic peptide	606	4	1.0	8	18	AAK2516.4	Hepatitis C virus
534	4	1.0	8	14	AAK6511.2	Hepatitis C virus	607	4	1.0	8	18	AAK2492.6	Hepatitis C virus
535	4	1.0	8	14	AAK6122.1	Hepatitis C virus	608	4	1.0	8	18	AAK2419.5	Hepatitis C virus
536	4	1.0	8	14	AAK6122.1	Hepatitis C virus	609	4	1.0	8	18	AAK2419.5	Hepatitis C virus
537	4	1.0	8	14	AAK6123.4	Hepatitis C virus	610	4	1.0	8	18	AAK2419.5	Hepatitis C virus
538	4	1.0	8	14	AAK6124.3	Hepatitis C virus	611	4	1.0	8	18	AAK2419.5	Hepatitis C virus
539	4	1.0	8	14	AAK6125.5	Hepatitis C virus	612	4	1.0	8	18	AAK2419.5	Hepatitis C virus
540	4	1.0	8	14	AAK4192.7	Peptide treatment of	613	4	1.0	8	19	AAK2086.5	Hepatitis C virus
541	4	1.0	8	14	AAK5404.2	Control peptide (A	614	4	1.0	8	19	AAK2074.5	Hepatitis C virus
542	4	1.0	8	15	AAK5538.3	Ion channel-formin	615	4	1.0	8	19	AAK4536.3	Hepatitis C virus
543	4	1.0	8	15	AAK4561.4	Ion channel-formin	616	4	1.0	8	19	AAK7428.3	Hepatitis C virus
544	4	1.0	8	15	AAK4580.5	Antibodies in type	617	4	1.0	8	19	AAK7429.1	Hepatitis C virus
545	4	1.0	8	15	AAK4580.6	Antibodies in type	618	4	1.0	8	19	AAK6651.9	Hepatitis C virus
546	4	1.0	8	15	AAK4580.7	Antibodies in type	619	4	1.0	8	19	AAK7916.8	Hepatitis C virus
547	4	1.0	8	15	AAK4580.8	Antibodies in type	620	4	1.0	8	19	AAK4286.5	Hepatitis C virus
548	4	1.0	8	15	AAK4580.9	Antibodies in type	621	4	1.0	8	19	AAK6286.6	Hepatitis C virus
549	4	1.0	8	15	AAK4581.1	Antibodies in type	622	4	1.0	8	19	AAK6103.5	Hepatitis C virus
550	4	1.0	8	15	AAK4581.2	Antibodies in type	623	4	1.0	8	19	AAK6144.4	Hepatitis C virus
551	4	1.0	8	15	AAK4581.3	Antibodies in type	624	4	1.0	8	19	AAK6147.7	Hepatitis C virus
552	4	1.0	8	15	AAK4581.4	Antibodies in type	625	4	1.0	8	19	AAK6176.4	Hepatitis C virus
553	4	1.0	8	15	AAK4581.7	Antibodies in type	626	4	1.0	8	19	AAK6192.2	Hepatitis C virus
554	4	1.0	8	15	AAK4581.8	Antibodies in type	627	4	1.0	8	19	AAK6195.7	Hepatitis C virus
555	4	1.0	8	15	AAK4581.9	Antibodies in type	628	4	1.0	8	19	AAK6195.7	Hepatitis C virus
556	4	1.0	8	15	AAK4582.0	Antibodies in type	629	4	1.0	8	19	AAK6292.4	Hepatitis C virus
557	4	1.0	8	15	AAK4582.2	Antibodies in type	630	4	1.0	8	19	AAK6292.4	Hepatitis C virus
558	4	1.0	8	15	AAK4582.3	Antibodies in type	631	4	1.0	8	19	AAK6292.4	Hepatitis C virus
559	4	1.0	8	15	AAK4582.7	Antibodies in type	632	4	1.0	8	19	AAK6292.4	Hepatitis C virus
560	4	1.0	8	15	AAK4583.0	Antibodies in type	633	4	1.0	8	19	AAK6292.4	Hepatitis C virus
561	4	1.0	8	15	AAK6955.9	Antibodies in type	634	4	1.0	8	19	AAK6292.4	Hepatitis C virus
562	4	1.0	8	15	AAK6955.9	Antibodies in type	635	4	1.0	8	19	AAK6292.4	Hepatitis C virus
563	4	1.0	8	15	AAK6955.9	Antibodies in type	636	4	1.0	8	19	AAK6292.4	Hepatitis C virus
564	4	1.0	8	15	AAK6955.9	Antibodies in type	637	4	1.0	8	19	AAK6292.4	Hepatitis C virus
565	4	1.0	8	15	AAK6955.9	Antibodies in type	638	4	1.0	8	19	AAK6292.4	Hepatitis C virus
566	4	1.0	8	15	AAK6955.9	Antibodies in type	639	4	1.0	8	19	AAK6292.4	Hepatitis C virus
567	4	1.0	8	15	AAK6955.9	Antibodies in type	640	4	1.0	8	19	AAK6292.4	Hepatitis C virus
568	4	1.0	8	15	AAK6955.9	Antibodies in type	641	4	1.0	8	19	AAK6292.4	Hepatitis C virus
569	4	1.0	8	15	AAK6955.9	Antibodies in type	642	4	1.0	8	19	AAK6292.4	Hepatitis C virus
570	4	1.0	8	15	AAK6955.9	Antibodies in type	643	4	1.0	8	19	AAK6292.4	Hepatitis C virus
571	4	1.0	8	15	AAK6955.9	Antibodies in type	644	4	1.0	8	19	AAK6292.4	Hepatitis C virus
572	4	1.0	8	15	AAK6955.9	Antibodies in type	645	4	1.0	8	19	AAK6292.4	Hepatitis C virus
573	4	1.0	8	15	AAK6955.9	Antibodies in type	646	4	1.0	8	19	AAK6292.4	Hepatitis C virus
574	4	1.0	8	15	AAK6955.9	Antibodies in type	647	4	1.0	8	19	AAK6292.4	Hepatitis C virus
575	4	1.0	8	15	AAK6955.9	Antibodies in type	648	4	1.0	8	19	AAK6292.4	Hepatitis C virus
576	4	1.0	8	15	AAK6955.9	Antibodies in type	649	4	1.0	8	19	AAK6292.4	Hepatitis C virus
577	4	1.0	8	15	AAK6955.9	Antibodies in type	650	4	1.0	8	19	AAK6292.4	Hepatitis C virus
578	4	1.0	8	15	AAK6955.9	Antibodies in type	651	4	1.0	8	19	AAK6292.4	Hepatitis C virus
579	4	1.0	8	15	AAK6955.9	Antibodies in type	652	4	1.0	8	19	AAK6292.4	Hepatitis C virus
580	4	1.0	8	15	AAK6955.9	Antibodies in type	653	4	1.0	8	19	AAK6292.4	Hepatitis C virus
581	4	1.0	8	15	AAK6955.9	Antibodies in type	654	4	1.0	8	19	AAK6292.4	Hepatitis C virus
582	4	1.0	8	15	AAK6955.9	Antibodies in type	655	4	1.0	8	19	AAK6292.4	Hepatitis C virus
583	4	1.0	8	15	AAK6955.9	Antibodies in type	656	4	1.0	8	19	AAK6292.4	Hepatitis C virus
584	4	1.0	8	15	AAK6955.9	Antibodies in type	657	4	1.0	8	19	AAK6292.4	Hepatitis C virus
585	4	1.0	8	15	AAK6955.9	Antibodies in type	658	4	1.0	8	19	AAK6292.4	Hepatitis C virus
586	4	1.0	8	15	AAK6955.9	Antibodies in type	659	4	1.0	8	19	AAK6292.4	Hepatitis C virus
587	4	1.0	8	15	AAK6955.9	Antibodies in type	660	4	1.0	8	19	AAK6292.4	Hepatitis C virus
588	4	1.0	8	15	AAK6955.9	Antibodies in type	661	4	1.0	8	19	AAK6292.4	Hepatitis C virus
589	4	1.0	8	15	AAK6955.9	Antibodies in type	662	4	1.0	8	19	AAK6292.4	Hepatitis C virus
590	4	1.0	8	15	AAK6955.9	Antibodies in type	663	4	1.0	8	19	AAK6292.4	Hepatitis C virus
591	4	1.0	8	15	AAK6955.9	Antibodies in type	664	4	1.0	8	19	AAK6292.4	Hepatitis C virus
592	4	1.0	8	15	AAK6955.9	Antibodies in type	665	4	1.0	8	19	AAK6292.4	Hepatitis C virus
593	4	1.0	8	15	AAK6955.9	Antibodies in type	666	4	1.0	8	19	AAK6292.4	Hepatitis C virus
594	4	1.0	8	15	AAK6955.9	Antibodies in type	667	4	1.0	8	19	AAK6292.4	Hepatitis C virus
595	4	1.0	8	15	AAK6955.9	Antibodies in type	668	4	1.0	8	19	AAK6292.4	Hepatitis C virus

[illegible]

APPENDIX

[illegible]

Query Match 1.7%; Score 7; DB 10; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6,4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 197 CGSCWA 292
 1111111
 Db 2 CGSCWA 8

RESULT 3

AAP90983
 ID AAP90983 standard; protein; 7 AA.

XX
 AC AAP90983;

XX 06-JUN-1990 (first entry)

DE Binding receptor with selectivity for a target ligand. Inverse by
 DE an article for inactivating toxic materials.

XX Toxic material: insecticide; organophosphorus cpds.; nerve poison;
 XX pesticide; decontaminant; military.

XX W08902920-A.

XX 06-APR-1989.

XX 04-OCT-1988; 88W0-0503422.

XX 05-OCT-1987; 87W0-0105412

XX (LITT) LITTLE AD INC.

XX 1aylor RF;

XX WPI: 1989-114395/15

PT Article for inactivating toxic materials, eg organophosphorus cpds.,
 PT comprises solid carrier bearing target ligand binding receptor

PT and ligand-degrading receptor, pref. enzyme

PS Claim 7b: page 42; 57pp; English.

XX The article for inactivating a toxic material comprises a solid carrier
 CC bearing a first receptor which binds the target ligand and a second

CC receptor which degrades the target ligand. This synthetic peptide is
 CC a preferred first receptor. The article may be used for covering

CC surfaces to protect or decontaminate the surface. The article is esp. for
 CC degrading toxic organic cpds., esp. organophosphorus cpds., (eg

CC pesticides and nerve poisons, bacteria and viruses, in environmental,
 CC chemical, military and industrial settings.

XX Sequence 7 AA:

Query Match 1.5%; Score 6; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6,4e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWA 292
 1111111

Db 2 CGSCWA 7

RESULT 4

AAG94916
 ID AAG94916 standard; Peptide; 10 AA.

XX AAG94916;

XX 18-SEP-2001 (first entry)

DE human complementary peptide, SEQ ID NO: 1110.

XX

KW human, complementary peptide, ligand, drug discovery, drug design.
 XX Homo sapiens.
 OS
 XX W-200142277 A2.
 XX 14-JUN-2001.
 PD 14-JUN-2001.

XX 14-DEC-2000; 2000W0-004776.

XX 14-DEC-1999; 99GB-0029464.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Houl JR;

XX WPI: 2001-408419/44.

PT A set of peptide ligands consisting of specific complementary peptides
 PT to proteins encoded by genes of the human genome, useful in an assay

PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs.

XX Example 4, Page 198, 64pp; English.

XX The invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides

CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead

CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA:

QY 168 SESEE 174
 1111111

Db 1 SESEE 6

Query Match 1.5%; Score 6; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 86;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 SESEE 174
 1111111

Db 1 SESEE 6

RESULT 5
 AAM05254
 ID AAM05254 standard; peptide; 5 AA.

XX AAM05254.

XX 02-JUN-1997 (first entry)

DE Cyclic pentapeptide #11 used in LH-RH receptor antagonist.

XX LH-RH releasing hormone receptor; LH-RH antagonist;
 KW sex hormone, human, anterior pituitary, cancer, prostate, uterus;

KW breast, pituitary, prostate, anterior pituitary, prostate, uterus;
 KW breast, pituitary, prostate, anterior pituitary, prostate, uterus;

KW multiocular ovarian syndrome; comedo; pregnancy; contraceptive;
 KW ovarian, muscular cycle, cystitis, wart, genital; streptococcus; spawling;

KW 11sh; testosterone; superantigen; leuporelin acetate.

XX Key Location/Qualifiers

FT Modified-site 1 /note "joined via a peptide linkage to leu5"

FT Modified-site 2 /note "1-azetidine-4-phenylpiperazine, b-form
 FT residue"

FT Misc-difference 4 /note "b-form residue"

FT Modified-site 5 /note "joined via a peptide linkage to lys1"

OS Synthetic.

SO Sequence 6 AA:

Query Match 1.28; Score 5; DB 19; Length 6;
Best Local Similarity 100.0%; Prod. No. 6.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCW 201
11111
DB 2 CGSCW 6

RESULT 8

AAR10048

ID AAR10048 standard; peptide: 6 AA

AC AAR10048:

DT 15-MAR-1991 (first entry)

DE Hypotensive oligopeptide.

XX Oral hypotensive; fig.

KW Ficus carica.

CS JPO2282394 A.

XX 19-NOV-1990.

PD 24-APR-1989, 89JP 0104265.

XX 24-APR-1989; 89JP-0104265.

PR (AGEN) AGENCY OF IND SCI TECH.

PA (SHOS) SHOWA SANGYO KK.

DP WPI: 1991-004496/01.

PT New oligopeptide hypotensive drug - based on alanine, valine,

XX asparagine, proline, isoleucine and arginine

PS Claim 1; Page 843; 9pp; Japanese.

CC Peptides may be derived from extract of fig. and are hypotensive

XX agents.

SO Sequence 6 AA:

Query Match 1.28; Score 5; DB 12; Length 6;
Best Local Similarity 100.0%; Prod. No. 6.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 188 VNP1K 192
11111
DB 2 VNP1K 6

RESULT 9

AAR57003

ID AAR57003 standard; peptide: 6 AA.

AC AAR57003:

DT 15-FEB-1995 (first entry)

DE N-terminal fragment of human histo-blood group A transferase.

XX Blood group determinant antigen; erythrocyte oligosaccharide;

KW glycoconjugate; glycosphingolipid; glycoprotein; glycosyltransferase;

XX transferase.

OS Homo sapiens.

XX

PN US5126857-A.

XX 05-JUN-1994.

PP 31 AUG 1989; 89US-0402695.

XX 31-AUG-1989; 89US-0402695.

PR 29-AUG 1991; 91US-0752101.

XX (HOMO-) BLOMEMBRANE INST.

PA Clausen H, Hakamori S, White T, Yamamoto F;

XX WPI: 1994-217996/26.

PT Isolated DNA molecules - encode human histo-blood groups A, B,

XX and O-glycosyltransferases

PS Example 2; Column 37; 64pp; English.

XX The histo-blood group ABO determinants are major allopathic antigens

CC in both erythrocytes and tissues of humans. They generally

CC constitute peripheral parts of the oligosaccharide chains of

CC glycoconjugates (e.g. linked to lipids (glycosphingolipids) or to

CC proteins (glycoproteins)). It was proposed that the A and B

CC phenotypes were associated with glycosyltransferases that converted

CC the H substance associated with the O phenotype to A and B

CC respectively, through the addition of specific N-acetylglucosamine

CC or alpha-1-3-galactosyl residues to the H and from fucose to the

CC beta-1-6. Hence, the primary products of the histo-blood group A

CC and B genes are the respective glycosyltransferases. This is a

CC fragment of the A group transferase. See also AAR57003 R57019.

XX Sequence 6 AA:

Query Match 1.28; Score 5; DB 15; Length 6;
Best Local Similarity 100.0%; Prod. No. 6.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 49 VLVVT 54
11111
DB 2 VLVVT 6

RESULT 10

AAW78339

ID AAW78339 standard; peptide: 6 AA.

AC AAW78339:

DT 27-APR-1999 (first entry)

DE Randomly encoded peptide protein fragment #13.

XX Expression; epitope; recombinant vector; structural gene; heterologous

XX identification.

XX Unidentified.

OS US5866463-A.

XX 02-FEB-1999.

XX 28-FEB-1991; 91US-0662764.

XX 28-FEB-1991; 91US-0662764.

XX 28-AUG-1985; 85US-0770390.

XX 26-MAY 1989; 89US-020158.

XX (FIBR?) FIBRININ G.

XX Pivcznik G.

XX

CC causes the intracellular accumulation of esters of cholesterol and the
CC formation of foam cells, with subsequent development of the
CC atherosclerotic plaque. The apo A IV-derived peptides can be used for
CC inhibiting lipid oxidation. The peptides can be used for treating
CC conditions associated with lipid oxidation. For example, they can be used
CC for treating or inhibiting the progression of atherosclerosis. They can
CC also be used for preventing oxidation in lipid-containing foods,
CC lipid-containing pharmaceuticals or cosmetic or dermatological
CC compositions. As these peptides comprise specific portions of the native
CC apo A-IV protein, there should be no immunogenicity problems with their
CC administration to humans.

XX
XX Sequence 7 AA:

Query Match 1.2%; Score 5; DB 20; Length 7;
Best local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 GLAFQ 251
IIIIII
DB 3 GLAFQ 7

RESULT 13

AAE10490
ID AAE10490 standard; peptide; 7 AA.

XX
XX AAE10490:

DT 10-DEC-2001 (first entry)

XX Humanised Ab clone 10 high potency modified light chain CDR, CDR-L2.

XX Human, light chain, respiratory syncytial virus infection, Alveolar,
XX parainfluenza virus, therapy, high potency antibody drug, cocaine,
XX cancer cell, toxic substance, complementarity determining region, Lf L2
XX
XX Homo sapiens.
XX Synthetic.

XX
XX Key Location/Qualifiers

FT Misc-difference 3 /note= "Wild type Ser substituted with Arg"

FT Misc-difference 4 /note= "Wild type Lys substituted with Gly"

FT Misc-difference 6 /note= "Wild type Ala substituted with Pro"

XX
XX WO200164751-A2.

XX
XX PD 07-SEP-2001.

XX
XX 01-MAR-2001; 2001WO-US06815.

XX
XX 01-MAR-2000; 2000MS-0186252

XX
XX (MED1-) MEDIMUNE INC.

PI Young JF, Koenig S, Johnson LS, Huse WB, Wu H, Watkins JJ

XX
XX WPI: 2001-582150/55.

XX
XX High potency monoclonal antibody, useful for preventing and treating

PT diseases caused or caused by viruses, especially respiratory syncytial

PT virus and parainfluenza virus, has high kinetic association rate.

PT constant.

XX
XX Claim 22, Page 33; 98pp; English.

XX
XX The invention relates to a high potency antibody, including its
XX immunologically active portions, fragments and segments other than
XX vitamin. The antibody has increased potency, high rate constant for
XX antibody-antigen complex formation and high affinity for any desired

CC action. The high potency antibody is also useful for reducing or
CC neutralizing the effects of addictive drugs, such as cocaine. The high
CC potency has specificity for antigenic determinants found on microbes
CC such as viruses, bacteria or fungi, antigens found on cancer cells and
CC toxic substances or product of toxic substances. The high potency
CC antibody is useful for preventing or treating a disease caused by a
CC virus such as respiratory syncytial virus (RSV) and parainfluenza
CC virus (PIV). The present sequence is humanised high potency modified
CC antibody light chain variable region complementarity determining region
CC (CDR).

XX
XX Sequence 7 AA:

Query Match 1.2%; Score 5; DB 22; Length 7;
Best local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 RGLFS 223
IIIIII
DB 3 RGLFS 7

RESULT 14

AAK43487
ID AAK43487 standard; peptide; 8 AA.

XX
XX AAK43487:

DT 12 MAY 1994 (first entry)

XX Ro/SSA epitope 484.

XX linearity: 1-16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1030, 1031, 1032, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071, 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081, 1082, 1083, 1084, 1085, 1086, 1087, 1088, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1096, 1097, 1098, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110, 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120, 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139, 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148, 1149, 1150, 1151, 1152, 1153, 1154, 1155, 1156, 1157, 1158, 1159, 1160, 1161, 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1170, 1171, 1172, 1173, 1174, 1175, 1176, 1177, 1178, 1179, 1180, 1181, 1182, 1183, 1184, 1185, 1186, 1187, 1188, 1189, 1190, 1191, 1192, 1193, 1194, 1195, 1196, 1197, 1198, 1199, 1200, 1201, 1202, 1203, 1204, 1205, 1206, 1207, 1208, 1209, 1210, 1211, 1212, 1213, 1214, 1215, 1216, 1217, 1218, 1219, 1220, 1221, 1222, 1223, 1224, 1225, 1226, 1227, 1228, 1229, 1230, 1231, 1232, 1233, 1234, 1235, 1236, 1237, 1238, 1239, 1240, 1241, 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1250, 1251, 1252, 1253, 1254, 1255, 1256, 1257, 1258, 1259, 1260, 1261, 1262, 1263, 1264, 1265, 1266, 1267, 1268, 1269, 1270, 1271, 1272, 1273, 1274, 1275, 1276, 1277, 1278, 1279, 1280, 1281, 1282, 1283, 1284, 1285, 1286, 1287, 1288, 1289, 1290, 1291, 1292, 1293, 1294, 1295, 1296, 1297, 1298, 1299, 1300, 1301, 1302, 1303, 1304, 1305, 1306, 1307, 1308, 1309, 1310, 1311, 1312, 1313, 1314, 1315, 1316, 1317, 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330, 1331, 1332, 1333, 1334, 1335, 1336, 1337, 1338, 1339, 1340, 1341, 1342, 1343, 1344, 1345, 1346, 1347, 1348, 1349, 1350, 1351, 1352, 1353, 1354, 1355, 1356, 1357, 1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1366, 1367, 1368, 1369, 1370, 1371, 1372, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380, 1381, 1382, 1383, 1384, 1385, 1386, 1387, 1388, 1389, 1390, 1391, 1392, 1393, 1394, 1395, 1396, 1397, 1398, 1399, 1400, 1401, 1402, 1403, 1404, 1405, 1406, 1407, 1408, 1409, 1410, 1411, 1412, 1413, 1414, 1415, 1416, 1417, 1418, 1419, 1420, 1421, 1422, 1423, 1424, 1425, 1426, 1427, 1428, 1429, 1430, 1431, 1432, 1433, 1434, 1435, 1436, 1437, 1438, 1439, 1440, 1441, 1442, 1443, 1444, 1445, 1446, 1447, 1448, 1449, 1450, 1451, 1452, 1453, 1454, 1455, 1456, 1457, 1458, 1459, 1460, 1461, 1462, 1463, 1464, 1465, 1466, 1467, 1468, 1469, 1470, 1471, 1472, 1473, 1474, 1475, 1476, 1477, 1478, 1479, 1480, 1481, 1482, 1483, 1484, 1485, 1486, 1487, 1488, 1489, 1490, 1491, 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499, 1500, 1501, 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511, 1512, 1513, 1514, 1515, 1516, 1517, 1518, 1519, 1520, 1521, 1522, 1523, 1524, 1525, 1526, 1527, 1528, 1529, 1530, 1531, 1532, 1533, 1534, 1535, 1536, 1537, 1538, 1539, 1540, 1541, 1542, 1543, 1544, 1545, 1546, 1547, 1548, 1549, 1550, 1551, 1552, 1553, 1554, 1555, 1556, 1557, 1558, 1559, 1560, 1561, 1562, 1563, 1564, 1565, 1566, 1567, 1568, 1569, 1570, 1571, 1572, 1573, 1574, 1575, 1576, 1577, 1578, 1579, 1580, 1581, 1582, 1583, 1584, 1585, 1586, 1587, 1588, 1589, 1590, 1591, 1592, 1593, 1594, 1595, 1596, 1597, 1598, 1599, 1600, 1601, 1602, 1603, 1604, 1605, 1606, 1607, 1608, 1609, 1610, 1611, 1612, 1613, 1614, 1615, 1616, 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629, 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639, 1640, 1641, 1642, 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651, 1652, 1653, 1654, 1655, 1656, 1657, 1658, 1659, 1660, 1661, 1662, 1663, 1664, 1665, 1666, 1667, 1668, 1669, 1670, 1671, 1672, 1673, 1674, 1675, 1676, 1677, 1678, 1679, 1680, 1681, 1682, 1683, 1684, 1685, 1686, 1687, 1688, 1689, 1690, 1691, 1692, 1693, 1694, 1695, 1696, 1697, 1698, 1699, 1700, 1701, 1702, 1703, 1704, 1705, 1706, 1707, 1708, 1709, 1710, 1711, 1712, 1713, 1714, 1715, 1716, 1717, 1718, 1719, 1720, 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1730, 1731, 1732, 1733, 1734, 1735, 1736, 1737, 1738, 1739, 1740, 1741, 1742, 1743, 1744, 1745, 1746, 1747, 1748, 1749, 1750, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758, 1759, 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769, 1770, 1771, 1772, 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780, 1781, 1782, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1

PE 29-SEP-1993: 93JN-0340853.
 XX
 XX (MTP) MITSUBISHI PETROCHEMICAL CO. LTD.
 PA
 DR WPI, 1995-17251/22
 XX
 PT New polypeptide(s) having affinity for the human TSH receptor
 PT antibody - used in detection of the TSH antibody.
 XX
 PS Example 1, Page 20, 54pp; Japanese.
 XX
 CC Peptides with affinity to human TSH (thyroid stimulating hormone)
 CC receptor antibody are used for detection of the antibody. (see also
 CC AAR73201-592).
 CC
 XX
 SQ Sequence 8 AA:
 Query Match 1.2%; Score 5; DB 16; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6, 4+0+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q7 255 KKKYL 259
 Db 1 KKKYL 5
 RESULT 18
 AAB12832
 ID AAB12832 standard; peptide; 8 AA.
 XX
 AC AAB12832:
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Src family kinase GGLT35-1 p110b, a.k.a. SRC, SEQ. IT NO. 12.
 XX
 KW APP Binding region; protein kinase A, PKA, CDK2, v-Src, inhibitor;
 KW target validation; mutant; enzyme; identification; protein kinases;
 KW protein kinase inhibitor; tumour, arthritis, cystostatic; cardiac;
 KW vasodilator; antiarrhythmic; analgesic; hyperproliferation; ischaemia;
 KW cardiovascular disease; urogenital disease; pain.
 XX
 OS Undifferentiated.
 OS
 XX
 PN W0200042042-A2
 XX
 PD 20-MAY-2000
 XX
 PF 11-JAN-2000: 990900-090951
 XX
 PR 11-JAN-1999: 9908-0115340.
 PR 23-MAY-1999: 9908-0145422
 XX
 PA (UYPB-) UNIV PRINCETON
 XX
 PI Shokal KM:
 XX
 DR WPI: 2000-491047/43.
 XX
 PT New enzyme inhibitors, useful for treating e.g. tumors or arthritis,
 PT are specific for mutant enzymes, without effect on wild type enzyme,
 PT particularly protein kinases -
 XX
 PS Example 14; Page 81; 169pp; English.
 XX
 CC The present invention describes an inhibitor (A) that does not inhibit
 CC a wild-type enzyme (E1) but does inhibit the same activity of a
 CC corresponding mutant enzyme (E2), with E1 and E2 being functionally
 CC identical. The inhibitor is specifically a protein kinase inhibitor
 CC which can have cytostatic, cardiac, vasodilator, antiarrhythmic and
 CC analgesic activities (A), particularly directed against protein kinases
 CC (PK), are used to disrupt oncogenic transformation and inhibit
 CC phosphorylation by mutant PK and growth of cells expressing mutant PK.

CC they can be used to treat tumours, hyperproliferation, cardiovascular
 CC and urogenital diseases, ischaemia, arthritis or pain. (A) are also
 CC useful for studying enzymatic function. Also (A), and mutant kinases,
 CC are used in gene therapy by knockout of a wild-type PK and replacement
 CC with a mutant PK which can then be regulated (switched on and off) by
 CC administration of (A). The protein kinase inhibitors can be used to
 CC disrupt transformation of cells that express a mutant Src-family PK.
 CC (A) are very specific for mutant enzymes, without significant effect on
 CC other PK. The present sequence represents a src family kinase optimised
 CC peptide substrate, which is used in an example from the present
 CC invention.
 XX
 SQ Sequence 8 AA:
 Query Match 1.2%; Score 5; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6, 4+0+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q7 26 EFKRK 42
 Db 4 EFKRK 8
 RESULT 19
 AAB08212
 ID AAB08212 standard; peptide; 8 AA.
 XX
 AC AAB08212:
 XX
 DT 01-NOV-2001 (first entry)
 XX
 DE Peptide #111 from human ribonucleoprotein, 60 kD Ro/SSA.
 XX
 KW Systemic lupus erythematosus, SLE; vaccine; immune response; therapeutics;
 KW autoimmune disorder; immune-protein; human; Ro/SSA peptide.
 XX
 OS Homo sapiens.
 OS
 XX
 PN 05623252-B1.
 XX
 PD 15-MAY-2001.
 XX
 PF 30-NOV-1993: 9308-0160604.
 XX
 PR 31 JAN 1993: 9308-0472947.
 PR 31-JAN-1991: 9108-0648205.
 PR 13-APR 1992: 9208-0867619.
 XX
 PA (OKLA) OKLAHOMA HERITAGE MUSEUM.
 XX
 PI Harley JR, James JA, Scottfield RH:
 XX
 DR WPI: 2001-3325087/35.
 XX
 PT Generating systemic lupus erythematosus animal model by immunizing
 PT non-human animal with non-immunoglobulin peptide having amino acid
 PT sequence of self-antigen bound by autoantibody population in early
 PT stage of disease -
 XX
 PS Example 1; Column 17; 63pp; English.
 XX
 CC The patent discloses a specific method of generating an animal model
 CC of systemic lupus erythematosus (SLE), comprising immunising a non-
 CC human animal with non-immunoglobulin peptide which comprises an
 CC epitope immunoreactive with auto-antibody (AAb) from patient with
 CC SLE. The epitope includes a region of self antigen which is bound
 CC by AAb population present in early stage in patient with SLE. The
 CC method is used for generating an animal model of SLE. It is useful
 CC for screening therapeutics effective in treating autoimmune disorders.
 CC It is useful as a component in a diagnostic assay, as a therapeutic
 CC vaccine to block the AAbs produced, by eliciting immune response,
 CC and in research on the possible causes of the autoimmune diseases.
 CC The method is used to interrupt the course of an autoimmune response.

CC once autoimmunity against the autoantigen is established. The amino
 CC acid sequences are used to make agents for neutralising circulating
 CC antibodies or immobilised on substrates in extracorporeal devices for
 CC specific removal of Aabs. The reagents identified by using the method
 CC are useful in manufacturing and testing autoantigens. The method is
 CC useful as models for screening of compounds which induce autoimmunity,
 CC inhibit induction of autoimmunity, suppress autoimmunity, it is useful
 CC in diagnosis of autoimmunity and as therapeutics for the treatment of
 CC autoimmune disorders. The present sequence is a peptide from human ribo-
 CC nucleoprotein, 60 kb RNS/SSA. This peptide is used in the exemplification
 CC of the invention.

XX Sequence 8 AA:

Query Match 1.28; Score 5; DB 14; Length 3;
 Host Local Similarity 100.0%; Pred. No. 6,444,057;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY R0 PFYRK 84
 DB 1 REYRK 5

RESULT 22

AAK22735 standard; Protein; 9 AA.

XX AAK22735.

XX 27-AUG-1992 (first entry)

XX Non-A, Non-B hepatitis virus antibody binding peptide.

XX NANBH; immunoreactive; detection.

XX Synthetic.

XX EP479376-A.

XX 08-APR-1992.

XX 26-SEP-1991 91PF-0202499

XX 05-OCT-1990; 90EP-0202658;

XX 26 SEP 1991; 91PF 0202498;

XX (ALKU) AKZO NV.

XX Habets WJA;

XX WPI; 1992-116131/15.

XX New peptide(s) reactive with antibodies to hepatitis non-A, non-B
 PT virus - for detection of NANBH and its antibodies in body fluids

XX Example; Page 7; 9pp; English.

CC The sequence is that of a peptide which is immunoreactive with
 CC antibodies to hepatitis non-A, non-B (NANBH) virus; it was pred. by
 CC stepwise solid phase peptide synthesis. It can be used to detect
 CC NANBH antibodies in test fluids. It has improved immunochromatol
 CC reactivity, a high affinity to NANBH antibodies and is of a satis
 CC non-infectious origin. See also AAK22737 E02744.

XX Sequence 9 AA:

Query Match 1.28; Score 5; DB 14; Length 9;
 Host Local Similarity 100.0%; Pred. No. 6,444,057;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 KTTNS 122
 DB 2 KTTNS 6

RESULT 23
 AAK44292

XX AAK44292 standard; peptide; 9 AA.

XX 09-JUN-1994 (first entry)

XX Residues 264-272 of human p53.

XX Human p53; protein; Class II; altered diagnostic assay; detection;
 KW p53 protein-specific T-cell antibody; cancer; M17 HLA A2.1.

XX Homo sapiens.

XX W09424925 A.

XX 09-DIC-1993.

XX H-KAY 1993. 91W-0105102.

XX 26-MAY-1992; 92EP-0201510.

XX (COTLE) ECKENHIV LEIDEN.

XX (SEED-) SEED CAPITAL INVESTMENTS (SOT) BV.

XX Kast WM, Method CUM;

XX WPI; 1993-405730/50.

XX Peptide derived from p53 protein, used to treat and diagnose
 PT diseases involving over expression of p53 e.g. human cancers

XX Claim 4; Page 48; 46pp; English.

CC The sequences given in AAK44281-94 are peptides derived from human p53
 CC protein which have the ability to bind to a human class I molecule.
 CC These peptides may be used in a diagnostic test or assay to detect
 CC human p53 protein-specific T-cells or antibodies. They may also be
 CC used in the treatment of diseases such as human cancers showing p53
 CC protein overexpression. These peptides can bind human MHC class I
 CC allele HLA A2.1.

XX Sequence 9 AA:

Query Match 1.28; Score 5; DB 14; Length 9;
 Host Local Similarity 100.0%; Pred. No. 6,444,057;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 LVKNS 363
 DB 2 LVKNS 6

RESULT 24

AAK56446

XX AAK56446 standard; peptide; 9 AA.

XX AAK56446;

XX 22-MAY-1994 (first entry)

XX Peptide inhibitor of selectin binding.

XX P-selectin; stannate membrane protein; 140; 2MF 140; P-selectin;
 KW E-selectin; inhibit neutrophil binding; diagnostic; therapeutic.

XX Synthetic.

XX Key location/Qualifiers

XX Modified-site 9

SU Sequence 9 AA:

Query Match 1.2% Score 5; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6, 4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 238 ALECA 212
IIIIII
DB 3 ALECA 7

RESULT 31

ABR55935
ID ABR55935 standard; Peptide: 9 AA.

AC ABR55935;
DI 15-FEB-2002 (first entry)

DE Vascular dementia-associated protein isoform (VPI) 135.

KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
diagnosis; prognosis; gene therapy.

OS Homo sapiens.

XX W0200169261 A2.

PD 20-SEP-2001.

PF 14-MAR-2001; 2001W0-GH01106.

PR 15-MAR-2000; 2000GH-0006285.

PR 24-NOV-2000; 2000GB-0028724.

PR 28-NOV-2000; 2000GB-0024391.

PA (OXFORD) OXFORD GLYCO-SCIENCES UK LTD.

PI Herath HMC, Parakk RB, Kohitt C;

PT Screening, diagnosis of prognosis of vascular dementia (VD), useful for
determining stage of VD and monitoring the effect of VD therapy.

PT comprises analysing body fluid by 2-dimensional electrophoresis for
features correlated with VD -

PS Claim 6; Page 32; 151pp; English.

CC The invention relates to screening, diagnosis or prognosis of Vascular
Dementia (VD) in a subject comprising analysing body fluid from the
subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
features containing at least one chosen feature whose relative abundance
correlates with the presence, absence, stage or severity of VD or
predicts the onset or course of VD, especially detecting in a sample of
cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
protein isoforms (VPIs) (ABR55801-ABR56295) as fully defined in the
specification. Detecting VD-associated features and VPI is useful for the
screening, diagnosis or prognosis of VD, for determining the stage or
severity of VD, for identifying a subject at risk of VD or for
monitoring the effect of therapy administered to a subject having VD.
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
useful for the treatment of VD and for gene therapy.

SQ Sequence 9 AA:

Query Match 1.2% Score 5; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6, 4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 IALIN 23
IIIIII
DB 2 IALIN 6

RESULT 32

ABR56033
ID ABR56033 standard; Peptide: 9 AA.

AC ABR56033;
DI 15-FEB-2002 (first entry)

DE Vascular dementia-associated protein isoform (VPI) 293.

KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
diagnosis; prognosis; gene therapy.

OS Homo sapiens.

XX W0200169261 A2.

PD 20-SEP-2001.

PF 14-MAR-2001; 2001W0-GH01106.

PR 15-MAR-2000; 2000GH-0006285.

PR 24-NOV-2000; 2000GB-0028724.

PR 28-NOV-2000; 2000GB-0024391.

PA (OXFORD) OXFORD GLYCO-SCIENCES UK LTD.

PI Herath HMC, Parakk RB, Kohitt C;

PT Screening, diagnosis of prognosis of vascular dementia (VD), useful for
determining stage of VD and monitoring the effect of VD therapy.

PT comprises analysing body fluid by 2 dimensional electrophoresis for
features correlated with VD -

PS Claim 6; Page 34; 151pp; English.

CC The invention relates to screening, diagnosis or prognosis of Vascular
Dementia (VD) in a subject comprising analysing body fluid from the
subject by 2 dimensional (2-D) electrophoresis to generate a 2-D array of
features containing at least one chosen feature whose relative abundance
correlates with the presence, absence, stage or severity of VD or
predicts the onset or course of VD, especially detecting in a sample of
cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
protein isoforms (VPIs) (ABR55801-ABR56295) as fully defined in the
specification. Detecting VD-associated features and VPI is useful for the
screening, diagnosis or prognosis of VD, for determining the stage or
severity of VD, for identifying a subject at risk of VD or for
monitoring the effect of therapy administered to a subject having VD.
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
useful for the treatment of VD and for gene therapy.

SQ Sequence 9 AA:

Query Match 1.2% Score 5; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6, 4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 IALIN 24
IIIIII
DB 2 IALIN 6

RESULT 33

ABR52192
ID ABR52192 standard; Peptide: 9 AA.

AC ABR52192;

DI 08-FEB-2002 (first entry)

PT Interleukin 1-beta - can form inhibitor compsn. and epids. for
PT treatment of arthritis, autoimmune diseases, inflammation and
PT radiation damage; also for wound healing
XX
PS Example: Page 37, 69pp; English.

CC The sequence is that of a peptide corresponding to Ala12 to Ser121
CC of human precursor interleukin (IL)-1beta sequence. It contains
CC a modified (His115 -> Phe) IL-1beta protease cleavage site
CC (His115-Pro118) for IL-1beta protease. It shows a reactivity of
CC 0.50 relative to the original sequence.
XX

XX Sequence 10 AA:

Query Match 1.28; Score 5; DB 14; Length 10;
Best Local Similarity 100.0%; Prod. No. 8.6e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 329 VEDAP 333
DB 3 VEDAP 7

RESULT 46

AA61618 standard; peptide: 10 AA.

AA61618;

12-MAY-1995 (first entry)

Peptide fragment (1.0579) of HBV binds HLA-A2.1.

XX antigen: epitope; immunogenic target protein; PSA; HBV; HBV; HBV;
XX HTLV; plasma specific antigen; hepatitis B virus; Epstein Barr;
XX human immunodeficiency virus; human Papilloma virus; P53; c-ERB2;
XX MAGE-1; melanoma antigen-1; core antigen; surface antigen;
XX pharmacological composition; in vitro; ex vivo; therapeutic;
XX diagnostic; MHC class I molecule; major histocompatibility complex;
XX HLA-A2.1; 9mer; 10mer; human leukocyte antigen.

XX Hepatitis B virus strain adr.

XX W09420127-A.

XX 15-SEP-1994.

XX 04-MAR-1994; 94W0-0802354

XX 05-MAR-1994; 94W0-0027146.

XX 04-JUN-1993; 93US-0073205.

XX 29-NOV-1993; 93US-0159184.

XX (CYTEC) CYTEL CORP.

XX Grey HM, Kasl WM, Settle A, Sidney JJ

XX WPI: 1994-302678/37.

XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used
XX for treatment or prophylaxis of cancer, virus infection or
XX autoimmune diseases.

XX Example 5; Page 112; 18pp; English.

XX AAR59456-561565 are immunogenic 10mer peptides that contain a HLA-A2.1
XX binding motif. These peptides bind HLA-A2.1 and have a binding
XX affinity of at least 1e as compared to a reference peptide (AAR71293).
XX AAR61618 occurs at position 1454 in the HBV "X" protein. The peptides
XX of the invention can induce cytotoxic T lymphocytes which can react
XX with target cells. They can be used for the treatment or prophylaxis
XX of cancer, eg prostate cancer or lymphoma, etc.

XX Sequence 10 AA:

Query Match 1.28; Score 5; DB 15; Length 10;
Best Local Similarity 100.0%; Prod. No. 8.6e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 202 AVEAV 206
DB 6 AVEAV 10

RESULT 47

AA65747 standard; peptide: 10 AA.

AA65747;

16-OCT-1998 (first entry)

Fibronectin binding protein-derived peptide #91.

XX microbial surface components recognising adhesive matrix molecules;
XX MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
XX antibody; bacterial infection; antibiotic resistant strain.

XX Synthetic.

XX Staphylococcus aureus.

XX W0983189-A2.

XX 23-JUL-1998.

XX 21-JAN-1998; 98W0-0801222.

XX 21-JAN-1997; 97US-0036139.

XX (TEXA) UNIV TEXAS A & M SYSTEM.

XX Hwang M, Hwang E, Kim J, Kim M, Park M, Park J, Kim

XX Special P;

XX WPI: 1998-414816/35.

XX Antibody that binds to fibronectin binding protein, preventing its
XX binding to fibronectin - used to treat or prevent bacterial
XX infection, especially by staphylococci and streptococci

XX Example 9; Page 111; 20pp; English.

XX The invention relates to antibodies that bind to a fibronectin binding
XX domain or a fibronectin-binding protein, and inhibit binding of the
XX protein to a fibronectin. Also claimed are: (1) isolated peptides of a
XX fibronectin-binding protein that do not bind to fibronectin; (2) fusion
XX protein containing at least one peptide of a fibronectin-binding protein
XX linked to a second amino acid sequence; (3) nucleic acid encoding
XX the isolated peptide of (1). Antibodies, the isolated peptides of
XX (1) and the nucleic acids are all useful for immunisation (active or
XX passive) and (by inhibiting binding of bacteria to fibronectin) for
XX preventing or treating infection in humans or other animals.
XX particularly by staphylococci or streptococci, e.g. meningitis, otitis
XX media, pneumonia, endocarditis, mastitis, in cattle, abortion in horses
XX and many others. Since the antibodies block binding of bacteria, they
XX should be effective against antibiotic resistant strains, and may
XX replace antibiotic therapy or improve its effectiveness. Sequences
XX AAW57447 represent peptides which were synthesized upon the 34
XX motif of S. aureus fibronectin binding protein A.

Query Match 1.28; Score 5; DB 19; Length 10;
Best Local Similarity 100.0%; Prod. No. 8.6e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

01. 101NK 000
11 1.000000

02. 101NK 000

11 1.000000

03. 101NK 000

11 1.000000

04. 101NK 000 (101NK 000)

11 1.000000

05. 101NK 000

11 1.000000

06. 101NK 000

11 1.000000

07. 101NK 000

11 1.000000

08. 101NK 000

11 1.000000

09. 101NK 000

11 1.000000

10. 101NK 000

11 1.000000

11. 101NK 000

12. 101NK 000

13. 101NK 000

14. 101NK 000

15. 101NK 000

16. 101NK 000

17. 101NK 000

18. 101NK 000

19. 101NK 000

20. 101NK 000

21. 101NK 000

22. 101NK 000

23. 101NK 000

24. 101NK 000

25. 101NK 000

26. 101NK 000

27. 101NK 000

01. 101NK 000
11 1.000000

02. 101NK 000

11 1.000000

03. 101NK 000

11 1.000000

04. 101NK 000 (101NK 000)

11 1.000000

05. 101NK 000

11 1.000000

06. 101NK 000

11 1.000000

07. 101NK 000

11 1.000000

08. 101NK 000

11 1.000000

09. 101NK 000

11 1.000000

10. 101NK 000

11 1.000000

11. 101NK 000

12. 101NK 000

13. 101NK 000

14. 101NK 000

15. 101NK 000

16. 101NK 000

17. 101NK 000

18. 101NK 000

19. 101NK 000

20. 101NK 000

21. 101NK 000

22. 101NK 000

23. 101NK 000

24. 101NK 000

25. 101NK 000

26. 101NK 000

27. 101NK 000

01. 101NK 000
11 1.000000

02. 101NK 000

11 1.000000

03. 101NK 000

11 1.000000

04. 101NK 000 (101NK 000)

11 1.000000

05. 101NK 000

11 1.000000

06. 101NK 000

11 1.000000

07. 101NK 000

11 1.000000

08. 101NK 000

11 1.000000

09. 101NK 000

11 1.000000

10. 101NK 000

11 1.000000

11. 101NK 000

12. 101NK 000

13. 101NK 000

14. 101NK 000

15. 101NK 000

16. 101NK 000

17. 101NK 000

18. 101NK 000

19. 101NK 000

20. 101NK 000

21. 101NK 000

22. 101NK 000

23. 101NK 000

24. 101NK 000

25. 101NK 000

26. 101NK 000

27. 101NK 000


```

XX MAb 17C7 binding site identifying M. catarrhalis USPA1 epitope #18.
DE Moraxella catarrhalis "USPA1" (USPA1) (USPA1) (USPA1) (USPA1)
XX Moraxella catarrhalis "USPA1" (USPA1) (USPA1) (USPA1) (USPA1)
XX vaccine: otitis media; sinusitis; lower respiratory tract infection;
KW immunity enhancer; immunoassay reagent; epitope; MAb 17C7.
XX Synthetic.
OS Moraxella catarrhalis.
XX W09828333-A2.
XX 02-JUL-1998.
XX 19-DEC-1997; 97WO-05239410.
XX 20-DEC-1996; 96US-0034508.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Aebi C, Coppe ID, Fiske MJ, Fiedler R, Hansel H.
XX Maciver I.
XX WPI: 1998-377595/32.
XX New peptide(s) containing the core epitope of Moraxella catarrhalis
XX USPA1 proteins - useful in, e.g. vaccines to prevent or treat M.
XX catarrhalis infection, and antibodies for passive immunisation.
XX Example 8: Page 113; 237pp; English.
XX Sequences shown in AAM68292 to AAM68296 represent decriptides of
XX Moraxella catarrhalis USPA1 antigen. These were used to identify a
XX binding site for MAb 17C7 to which a decriptide epitope of USPA1 antigen
XX can bind to. Nucleic acid sequences encoding the USPA1 and A2 antigens
XX of M. catarrhalis isolates 0567, 0466, T7A31 and T7A17 can be used to
XX genetic vaccination. An antigenic composition of vaccine containing
XX antigenic peptides from USPA1 or USPA2 antigens are used to induce an
XX immune response in mammals against M. catarrhalis and can be used to
XX treat infections such as otitis media, sinusitis, lower respiratory
XX tract infections. They can also be used as immunity enhancers for other
XX bacterial parasites of viral origin, to raise antibodies and as
XX immunoassay reagents for detecting specific antibodies. The antibodies
XX are useful for passive immunisation and as immunoassay reagents.
XX Detection of the epitopic core sequence, by immunoassay or by PCR, is
XX used to diagnose infection. The USPA1 antigens encoding nucleic acid
XX sequences are also used to produce recombinant proteins and for
XX screening for potential anti-M. catarrhalis agents, while their fragments
XX are useful as diagnostic probes or primers or to isolate variant
XX sequences.
XX Sequence 10 AA:
SO
Query Match 1.2%; Score 5; DB 19; Length 10;
Best Local Similarity 100.0%; Pred. No. 8; 66/02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 NNVVE 9
DB 4 NNVVE 8

```

```

XX Synapsin containing complex; SC complex; screening; modulate;
XX secretory release; synaptic transmission; secretory proteins;
XX treatment; affective disorder; depressive manic-depressive disorder;
XX anxiety disorder; neurodegenerative disease; radiopharmaceutical;
XX hormonal imbalance; antigen processing.
XX Rattus sp.
XX Key Location/Qualifiers
XX Misc-difference 1 /note "not specified"
XX W09828419-A2.
XX 02-JUL-1998.
XX 19-DEC-1997; 97WO-05239410.
XX 20-DEC-1996; 96US-0034508.
XX (TEXA) UNIV TEXAS SYSTEM.
XX Schellier RH.
XX WPI: 1998-377650/32.
XX New isolated vesicle secretion associated polypeptide(s) used to
XX develop products for treating e.g. affective disorder,
XX neurodegenerative disease, hormone imbalances, immune system
XX disorders or tumours.
XX Example 6: Page 30; 144pp; English.
XX AAM6929-35 represent peptide fragments of a p102 protein, which is
XX present in the secretion associated 173 (SA 173) and SA 174 proteins
XX from the SA 173 cells, which binds a secretory protein (SP) and
XX The SA 173 polypeptides and nucleotide sequences encoding them can be
XX used for screening for compounds which modulate vesicular release
XX involved in synaptic transmission and other secretory processes.
XX Compounds which enhance binding between the SA 173 and SP complexes may
XX be used to treat an affective disorder such as depression,
XX manic depressive disorder and anxiety disorders, or a neurodegenerative
XX disease such as Parkinson's disease or Huntington's disease. Compounds
XX which inhibit binding between the SA 173 and SP complexes may be used to
XX treat a disorder of thought, such as schizophrenia, or for anaesthesia.
XX The compounds can also be used to intervene in the endocrine system for
XX treatment of hormonal imbalances, the immune system for intervention in
XX antigen processing, secreted immunomodulators, and viral processing, as
XX well as anti-tumour applications, such as regulation of membrane
XX trafficking during rapid cell division.
XX Sequence 10 AA:
SO
Query Match 1.2%; Score 5; DB 19; Length 10;
Best Local Similarity 100.0%; Pred. No. 8; 66/02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 280 ENYIE 284
DB 4 ENYIE 7

```


CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.

XX Sequence 10 AA:

Query Match 1.2%; Score 5; DB 22; Length 10;

Best local Similarity 100.0%; Prod. No. 8.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ASESSE 171

Db 5 ASESSE 9

RESULT 45

AAG83680
ID AAG83680 standard; Peptide; 10 AA.

XX AAG83680;

DT 11-SEP-2001 (first entry)

DE Arabidopsis thaliana peptide ligand #120.

KW Plant; peptide pesticide; peptide herbicide; agricultural research.

OS Arabidopsis thaliana.

PN W0200142279-A2.

PD 14-JUN-2001.

PE 13-DEC-2000; 2000WO-GH04781

PR 13-DEC-1999; 99GB-0029469.

XX (PROT-) PROTEOM LTD.

PA Roberts GW, Heal JR;

DP WPI; 2001-201429/40

PT A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes.

PS Example 4; Page 75; 201pp; English.

CC The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.

XX Sequence 10 AA:

Query Match 1.2%; Score 5; DB 22; Length 10;

Best local Similarity 100.0%; Prod. No. 8.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ASESSE 171

Db 6 ASESSE 10

RESULT 46

AAG83952
ID AAG83952 standard; Peptide; 10 AA.

XX

AC AAG83952;

DT 11-SEP-2001 (first entry)

DE Arabidopsis thaliana peptide ligand #592.

KW Plant; peptide pesticide; peptide herbicide; agricultural research.

OS Arabidopsis thaliana.

PN W0200142279-A2.

PD 14-JUN-2001.

PE 13-DEC-2000; 2000WO-GH04781

PR 13-DEC-1999; 99GB-0029469.

XX (PROT-) PROTEOM LTD.

PA Roberts GW, Heal JR;

DP WPI; 2001-201429/40.

PT A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes.

PS Example 4; Page 114; 201pp; English.

CC The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.

XX Sequence 10 AA:

Query Match 1.2%; Score 5; DB 22; Length 10;

Best local Similarity 100.0%; Prod. No. 8.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 SAVAA 208

Db 1 SAVAA 5

RESULT 47

AAG83954
ID AAG83954 standard; Peptide; 10 AA.

XX AAG83954;

DT 11-SEP-2001 (first entry)

DE Arabidopsis thaliana peptide ligand #594.

KW Plant; peptide pesticide; peptide herbicide; agricultural research.

OS Arabidopsis thaliana.

PN W0200142279-A2.

PD 14-JUN-2001.

PE 13-DEC-2000; 2000WO-GH04781.

PR 13-DEC-1999; 99GB-0029469.

XX (PROT-) PROTEOM LTD.

KW Saccharomyces cerevisiae, Complementary peptide, peptide identification;
 KW drug discovery; drug design.

XX Saccharomyces cerevisiae.

XX WO200142270-A1.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-0804773.

XX 13-DEC-1999; 99GH-0029471.

XX (PROT-) PROTHEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis of protein and
 PT nucleotide sequence databases, useful in drug design.

XX Example 3; Page 186; 488pp; English.

CC The invention relates to the identification of complementary peptides
 CC by analysis of protein and nucleotide sequence databases from higher
 CC eukaryotic genomes, excluding human and plants. The specific
 CC complementary peptides interact with their relevant target proteins
 CC encoded in the eukaryotic genome. The peptides may be used as reagents
 CC and drugs for drug discovery and as lead ligands for drug design and
 CC development. The present sequence is a complementary peptide from
 CC Saccharomyces cerevisiae.

XX Sequence 10 AA;

SO Query Match

1.2%; Score 5; DB 22; Length 10;

Best Local Similarity 100.0%; Pred No R 60.02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 ESEEE 173

DB 3 ESEEE 7

Search completed October 9, 2002, 12:43:30
 Job time : 40 secs



249	2	0.7	10	2	B86897	T cell receptor α	422	0.5	4	2	P16677	T cell receptor β
250	2	0.7	10	2	A61354	cardiac mediating	423	0.5	4	2	P16651	T cell receptor β
251	3	0.7	16	2	P80660	lysosome-associated	424	0.5	4	2	P16697	T cell receptor β
252	3	0.7	10	2	C40572	T cell receptor β	425	0.5	4	2	P16706	T cell receptor β
253	3	0.7	10	2	P10927	T cell receptor β	426	0.5	4	2	P10721	T cell receptor β
254	3	0.7	10	2	P10925	T cell receptor β	427	0.5	4	2	P16566	T cell receptor β
255	3	0.7	10	2	P10923	T cell receptor β	428	0.5	4	2	A60155	T cell receptor β
256	3	0.7	10	2	T17054	cytochrome c α 11a	429	0.5	4	2	C47562	cytochrome c α 11a
257	3	0.7	10	2	T12308	cytochrome c α 11a	430	0.5	4	2	A62209	cytochrome c α 11a
258	3	0.7	10	2	T12316	cytochrome c α 11a	431	0.5	4	2	SEC348	cytochrome c α 11a
259	3	0.7	10	2	T12321	cytochrome c α 11a	432	0.5	5	1	BRCH4	cytochrome c α 11a
260	3	0.7	10	2	S63789	neuropeptide Y α II	433	0.5	5	2	J80862	neuropeptide Y α II
261	3	0.5	3	3	GKR0	growth-modulating	434	0.5	5	2	J80860	growth-modulating
262	3	0.5	3	3	A60894	keratin α -keratin	435	0.5	5	2	A62516	keratin α -keratin
263	3	0.5	3	3	A23751	serpin α 1	436	0.5	5	2	C23761	serpin α 1
264	3	0.5	3	3	P23751	serpin α 1	437	0.5	5	2	A60890	serpin α 1
265	3	0.5	3	3	A63802	lysozyme α -lysozyme	438	0.5	5	2	C41235	lysozyme α -lysozyme
266	3	0.5	3	3	A29566	lysozyme α -lysozyme	439	0.5	5	2	A11255	lysozyme α -lysozyme
267	3	0.5	3	3	A43391	lysozyme α -lysozyme	440	0.5	5	2	B41225	lysozyme α -lysozyme
268	3	0.5	3	3	P23192	bradykinin potenti	441	0.5	5	2	I40702	bradykinin potenti
269	3	0.5	3	3	P23196	bradykinin potenti	442	0.5	5	2	I40694	bradykinin potenti
270	3	0.5	3	3	P10636	T cell receptor β	443	0.5	5	2	B47325	T cell receptor β
271	3	0.5	3	3	P10578	T cell receptor β	444	0.5	5	2	A82014	T cell receptor β
272	3	0.5	3	3	P10571	T cell receptor β	445	0.5	5	2	B42604	T cell receptor β
273	3	0.5	3	3	P10632	T cell receptor β	446	0.5	5	2	A41566	T cell receptor β
274	3	0.5	3	3	S63326	blood cell protein	447	0.5	5	2	I40459	blood cell protein
275	3	0.5	3	3	T13812	cytochrome c α 11a	448	0.5	5	2	P60274	cytochrome c α 11a
276	3	0.5	4	1	EXAA	anthropamide neur	449	0.5	5	2	P60274	anthropamide neur
277	3	0.5	4	1	A61147	phosphatidylserine	450	0.5	5	2	B26274	phosphatidylserine
278	3	0.5	4	1	A61147	phosphatidylserine	451	0.5	5	2	B26274	phosphatidylserine
279	3	0.5	4	1	ECRK	lysine melanocty	452	0.5	5	2	T10571	lysine melanocty
280	3	0.5	4	1	P10140	cardiac melanocty	453	0.5	5	2	T10571	cardiac melanocty
281	3	0.5	4	1	P10140	cardiac melanocty	454	0.5	5	2	A14892	cardiac melanocty
282	3	0.5	4	1	A61147	cardiac melanocty	455	0.5	5	2	P60324	cardiac melanocty
283	3	0.5	4	1	A61147	cardiac melanocty	456	0.5	5	2	P60324	cardiac melanocty
284	3	0.5	4	1	A61147	cardiac melanocty	457	0.5	5	2	P60324	cardiac melanocty
285	3	0.5	4	1	A61147	cardiac melanocty	458	0.5	5	2	P60324	cardiac melanocty
286	3	0.5	4	1	A61147	cardiac melanocty	459	0.5	5	2	P60324	cardiac melanocty
287	3	0.5	4	1	A61147	cardiac melanocty	460	0.5	5	2	P60324	cardiac melanocty
288	3	0.5	4	1	A61147	cardiac melanocty	461	0.5	5	2	P60324	cardiac melanocty
289	3	0.5	4	1	A61147	cardiac melanocty	462	0.5	5	2	P60324	cardiac melanocty
290	3	0.5	4	1	A61147	cardiac melanocty	463	0.5	5	2	P60324	cardiac melanocty
291	3	0.5	4	1	A61147	cardiac melanocty	464	0.5	5	2	P60324	cardiac melanocty
292	3	0.5	4	1	A61147	cardiac melanocty	465	0.5	5	2	P60324	cardiac melanocty
293	3	0.5	4	1	A61147	cardiac melanocty	466	0.5	5	2	P60324	cardiac melanocty
294	3	0.5	4	1	A61147	cardiac melanocty	467	0.5	5	2	P60324	cardiac melanocty
295	3	0.5	4	1	A61147	cardiac melanocty	468	0.5	5	2	P60324	cardiac melanocty
296	3	0.5	4	1	A61147	cardiac melanocty	469	0.5	5	2	P60324	cardiac melanocty
297	3	0.5	4	1	A61147	cardiac melanocty	470	0.5	5	2	P60324	cardiac melanocty
298	3	0.5	4	1	A61147	cardiac melanocty	471	0.5	5	2	P60324	cardiac melanocty
299	3	0.5	4	1	A61147	cardiac melanocty	472	0.5	5	2	P60324	cardiac melanocty
300	3	0.5	4	1	A61147	cardiac melanocty	473	0.5	5	2	P60324	cardiac melanocty
301	3	0.5	4	1	A61147	cardiac melanocty	474	0.5	5	2	P60324	cardiac melanocty
302	3	0.5	4	1	A61147	cardiac melanocty	475	0.5	5	2	P60324	cardiac melanocty
303	3	0.5	4	1	A61147	cardiac melanocty	476	0.5	5	2	P60324	cardiac melanocty
304	3	0.5	4	1	A61147	cardiac melanocty	477	0.5	5	2	P60324	cardiac melanocty
305	3	0.5	4	1	A61147	cardiac melanocty	478	0.5	5	2	P60324	cardiac melanocty
306	3	0.5	4	1	A61147	cardiac melanocty	479	0.5	5	2	P60324	cardiac melanocty
307	3	0.5	4	1	A61147	cardiac melanocty	480	0.5	5	2	P60324	cardiac melanocty
308	3	0.5	4	1	A61147	cardiac melanocty	481	0.5	5	2	P60324	cardiac melanocty
309	3	0.5	4	1	A61147	cardiac melanocty	482	0.5	5	2	P60324	cardiac melanocty
310	3	0.5	4	1	A61147	cardiac melanocty	483	0.5	5	2	P60324	cardiac melanocty
311	3	0.5	4	1	A61147	cardiac melanocty	484	0.5	5	2	P60324	cardiac melanocty
312	3	0.5	4	1	A61147	cardiac melanocty	485	0.5	5	2	P60324	cardiac melanocty
313	3	0.5	4	1	A61147	cardiac melanocty	486	0.5	5	2	P60324	cardiac melanocty
314	3	0.5	4	1	A61147	cardiac melanocty	487	0.5	5	2	P60324	cardiac melanocty
315	3	0.5	4	1	A61147	cardiac melanocty	488	0.5	5	2	P60324	cardiac melanocty
316	3	0.5	4	1	A61147	cardiac melanocty	489	0.5	5	2	P60324	cardiac melanocty
317	3	0.5	4	1	A61147	cardiac melanocty	490	0.5	5	2	P60324	cardiac melanocty
318	3	0.5	4	1	A61147	cardiac melanocty	491	0.5	5	2	P60324	cardiac melanocty
319	3	0.5	4	1	A61147	cardiac melanocty	492	0.5	5	2	P60324	cardiac melanocty
320	3	0.5	4	1	A61147	cardiac melanocty	493	0.5	5	2	P60324	cardiac melanocty
321	3	0.5	4	1	A61147	cardiac melanocty	494	0.5	5	2	P60324	cardiac melanocty

833	2	0.5	9	2	326850	lg heavy chain V r	906	0.5	10	2	160787	sperm-act-ivad f
834	2	0.5	9	2	326740	lg kappa chain c r	907	0.5	10	2	A60787	sperm-act-ivad f
835	2	0.5	9	2	A60729	1-ethylparao-cm	908	0.5	10	2	B60787	sperm-act-ivad f
836	2	0.5	9	2	173279	T cell receptor bc	909	0.5	10	2	A60788	sperm-act-ivad f
837	2	0.5	9	2	170562	T cell receptor bc	910	0.5	10	2	A60727	sperm-act-ivad f
838	2	0.5	9	2	167073	ubiquitin-cytocb	911	0.5	10	2	G60527	sperm-act-ivad f
839	2	0.5	9	2	167078	ubiquitin-48-74	912	0.5	10	2	B60527	sperm-act-ivad f
840	2	0.5	9	2	158350	pro-emp1 protein	913	0.5	10	2	B60527	sperm-act-ivad f
841	2	0.5	9	2	149406	lane-4a protein	914	0.5	10	2	B60527	sperm-act-ivad f
842	2	0.5	9	2	A42256	EpH4/2A/2A/2A	915	0.5	10	2	G60535	sperm-act-ivad f
843	2	0.5	9	2	578426	ribosomal protein	916	0.5	10	2	G60535	sperm-act-ivad f
844	2	0.5	9	2	152974	sciniol vesicle pr	917	0.5	10	2	B60589	sperm-act-ivad f
845	2	0.5	9	2	PH0942	w-cell receptor bc	918	0.5	10	2	A60589	sperm-act-ivad f
846	2	0.5	9	2	PH0943	T cell receptor bc	919	0.5	10	2	B60588	sperm-act-ivad f
847	2	0.5	9	2	PH0935	T cell receptor bc	920	0.5	10	2	160527	sperm-act-ivad f
848	2	0.5	9	2	PH0932	T cell receptor bc	921	0.5	10	2	160527	sperm-act-ivad f
849	2	0.5	9	2	PH0917	T-cell receptor bc	922	0.5	10	2	X6097	sperm-act-ivad f
850	2	0.5	9	2	PH0918	T cell receptor bc	923	0.5	10	2	X6097	sperm-act-ivad f
851	2	0.5	9	2	PH0921	T cell receptor bc	924	0.5	10	2	A60574	sperm-act-ivad f
852	2	0.5	9	2	578426	52.5K protein	925	0.5	10	2	A60574	sperm-act-ivad f
853	2	0.5	9	2	578426	52.5K protein	926	0.5	10	2	A60574	sperm-act-ivad f
854	2	0.5	9	2	556004	glucan 1,3 beta-4	927	0.5	10	2	A24857	sperm-act-ivad f
855	2	0.5	9	2	170927	for tachykinin f	928	0.5	10	2	578426	sperm-act-ivad f
856	2	0.5	9	2	157650	hemoglobin alpha p	929	0.5	10	2	S24186	sperm-act-ivad f
857	2	0.5	9	2	170924	hypocretin-1	930	0.5	10	2	578426	sperm-act-ivad f
858	2	0.5	9	2	515594	ort 1 para 5'-recl	931	0.5	10	2	578426	sperm-act-ivad f
859	2	0.5	9	2	515595	ort 2 para 5'-recl	932	0.5	10	2	578426	sperm-act-ivad f
860	2	0.5	10	1	XAV168	angiotensin-conver	933	0.5	10	2	B34143	sperm-act-ivad f
861	2	0.5	10	1	PH093	gonadoliborin - fi	934	0.5	10	2	A34143	sperm-act-ivad f
862	2	0.5	10	1	RHSHG	gonadoliborin - sh	935	0.5	10	2	D37497	sperm-act-ivad f
863	2	0.5	10	1	RHSHG	gonadoliborin - sh	936	0.5	10	2	A61617	sperm-act-ivad f
864	2	0.5	10	1	RHSHG	gonadoliborin - sh	937	0.5	10	2	B46454	sperm-act-ivad f
865	2	0.5	10	1	PH092	gonadoliborin - sh	938	0.5	10	2	D37497	sperm-act-ivad f
866	2	0.5	10	1	A61126	gonadoliborin - sf	939	0.5	10	2	551912	sperm-act-ivad f
867	2	0.5	10	1	RHMG5	gonadoliborin - sf	940	0.5	10	2	S09487	sperm-act-ivad f
868	2	0.5	10	1	PH091M	tachykinin 1 - mlt	941	0.5	10	2	A24192	sperm-act-ivad f
869	2	0.5	10	1	PH091M	tachykinin 1 - mlt	942	0.5	10	2	B34617	sperm-act-ivad f
870	2	0.5	10	1	SPONK	neurokinin K - B14	943	0.5	10	2	139202	sperm-act-ivad f
871	2	0.5	10	1	SPONK	neurokinin K - B14	944	0.5	10	2	S06964	sperm-act-ivad f
872	2	0.5	10	1	SPONK	neurokinin K - B14	945	0.5	10	2	S06964	sperm-act-ivad f
873	2	0.5	10	1	SPONK	neurokinin K - B14	946	0.5	10	2	S06964	sperm-act-ivad f
874	2	0.5	10	1	SPONK	neurokinin K - B14	947	0.5	10	2	S06964	sperm-act-ivad f
875	2	0.5	10	1	SPONK	neurokinin K - B14	948	0.5	10	2	S06964	sperm-act-ivad f
876	2	0.5	10	1	SPONK	neurokinin K - B14	949	0.5	10	2	S06964	sperm-act-ivad f
877	2	0.5	10	1	SPONK	neurokinin K - B14	950	0.5	10	2	S06964	sperm-act-ivad f
878	2	0.5	10	1	SPONK	neurokinin K - B14	951	0.5	10	2	S06964	sperm-act-ivad f
879	2	0.5	10	1	SPONK	neurokinin K - B14	952	0.5	10	2	S06964	sperm-act-ivad f
880	2	0.5	10	1	SPONK	neurokinin K - B14	953	0.5	10	2	S06964	sperm-act-ivad f
881	2	0.5	10	1	SPONK	neurokinin K - B14	954	0.5	10	2	S06964	sperm-act-ivad f
882	2	0.5	10	1	SPONK	neurokinin K - B14	955	0.5	10	2	S06964	sperm-act-ivad f
883	2	0.5	10	1	SPONK	neurokinin K - B14	956	0.5	10	2	S06964	sperm-act-ivad f
884	2	0.5	10	1	SPONK	neurokinin K - B14	957	0.5	10	2	S06964	sperm-act-ivad f
885	2	0.5	10	1	SPONK	neurokinin K - B14	958	0.5	10	2	S06964	sperm-act-ivad f
886	2	0.5	10	1	SPONK	neurokinin K - B14	959	0.5	10	2	S06964	sperm-act-ivad f
887	2	0.5	10	1	SPONK	neurokinin K - B14	960	0.5	10	2	S06964	sperm-act-ivad f
888	2	0.5	10	1	SPONK	neurokinin K - B14	961	0.5	10	2	S06964	sperm-act-ivad f
889	2	0.5	10	1	SPONK	neurokinin K - B14	962	0.5	10	2	S06964	sperm-act-ivad f
890	2	0.5	10	1	SPONK	neurokinin K - B14	963	0.5	10	2	S06964	sperm-act-ivad f
891	2	0.5	10	1	SPONK	neurokinin K - B14	964	0.5	10	2	S06964	sperm-act-ivad f
892	2	0.5	10	1	SPONK	neurokinin K - B14	965	0.5	10	2	S06964	sperm-act-ivad f
893	2	0.5	10	1	SPONK	neurokinin K - B14	966	0.5	10	2	S06964	sperm-act-ivad f
894	2	0.5	10	1	SPONK	neurokinin K - B14	967	0.5	10	2	S06964	sperm-act-ivad f
895	2	0.5	10	1	SPONK	neurokinin K - B14	968	0.5	10	2	S06964	sperm-act-ivad f
896	2	0.5	10	1	SPONK	neurokinin K - B14	969	0.5	10	2	S06964	sperm-act-ivad f
897	2	0.5	10	1	SPONK	neurokinin K - B14	970	0.5	10	2	S06964	sperm-act-ivad f
898	2	0.5	10	1	SPONK	neurokinin K - B14	971	0.5	10	2	S06964	sperm-act-ivad f
899	2	0.5	10	1	SPONK	neurokinin K - B14	972	0.5	10	2	S06964	sperm-act-ivad f
900	2	0.5	10	1	SPONK	neurokinin K - B14	973	0.5	10	2	S06964	sperm-act-ivad f
901	2	0.5	10	1	SPONK	neurokinin K - B14	974	0.5	10	2	S06964	sperm-act-ivad f
902	2	0.5	10	1	SPONK	neurokinin K - B14	975	0.5	10	2	S06964	sperm-act-ivad f
903	2	0.5	10	1	SPONK	neurokinin K - B14	976	0.5	10	2	S06964	sperm-act-ivad f
904	2	0.5	10	1	SPONK	neurokinin K - B14	977	0.5	10	2	S06964	sperm-act-ivad f
905	2	0.5	10	1	SPONK	neurokinin K - B14	978	0.5	10	2	S06964	sperm-act-ivad f

DB 1 MAR 3

RESULT 17

184439

protamine P1 - savannah baboon (fragment)

C:Species: Papio hamadryas doquiera (savannah baboon)

C:Date: 19-Mar-1997 #sequence_revision 07-Nov-1997 #text_change 21 Jul 2000

C:Accession: 184439

R:Queralt, R.; Oliva, R.

Gene 133, 197-204, 1993

A:Title: Identification of conserved potential regulatory sequences of the protamine and

A:Reference number: 137013, MIMD 44060818

A:Accession: 184439

A:Status: preliminary, translated from GH/PMGL/TERP3

A:Molecule type: DNA

A:Residues: 1-4 <RES>

A:Cross-references: EMBL:212147, NID:438134, FIDN:CAAT3131.1, FIL:4167315

Query Match

Best Local Similarity	0.78;	Score 3;	DB 2;	Length 4;
Best Local	Similarity 100.0%;	Prod No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 2;	Indels 3;
			Gaps 2;	

QY 143 MAR 145

DB 1 MAR 3

RESULT 18

A60521

glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)

M:Alternate names: glycogen phosphorylase b

C:Species: Liza ramada

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-May-2000

C:Accession: A60521

R:Bonamusa, L.; Baanante, I.V.

Comp Biochem Physiol B 95, 295-301, 1990

A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus

A:Reference number: A60521; MIMD:90227907

A:Accession: A60521

A:Molecule type: protein

A:Residues: 1-5 <RES>

C:Superfamily: phosphorylase

C:Keywords: glycogen phosphorylase, howy phosphorylase, phosphorylase

E3/kinning site: phosphate (Ser) (catalytic) (by phosphorylase kinase) #status: confirmed

Query Match	0.78;	Score 3;	DB 2;	Length 5;
Best Local Similarity	100.0%;	Prod No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY	311	ISV	313
DB	2	ISV	4

Query Match	0.78;	Score 3;	DB 2;	Length 5;
Best Local Similarity	100.0%;	Prod No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY 311 ISV 313

DB 2 ISV 4

RESULT 19

S70154

URF2 protein - Xanthomonas sp.

C:Species: Xanthomonas sp.

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Jul-1999

C:Accession: S70154

R:Khalid, O.Y.; Mitchell, S.F.; Raza, I.A.; Yousif, O.V.; Mubashir, S.V.; NIKI, S.V.

Mol. Microbiol. 17, 1189-1200, 1995

A:Title: Four genes, two ends, and a few regions are involved in transposition of Tn3933

A:Reference number: S70140; MIMD:96130850

A:Accession: S70154

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: EMBL:140585; NID:9710572; PIRN:AA049129; J:PIR:Q73690

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, May 1997

Query Match

Best Local Similarity	0.78;	Score 3;	DB 2;	Length 5;
Best Local	Similarity 100.0%;	Prod No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

Best Local Similarity	100.0%;	Prod. No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY 166 SAS 168

DB 2 SAS 4

RESULT 20

B31836

28K protein - *Escherichia coli* (fragment)C:Species: *Escherichia coli* (strain)

C:Date: 31-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 28-May-1999

C:Accession: B31836

R:Anderson, R.E.; Baumstark, B.R.; Bellini, W.J.

J. Bacteriol. 170, 4493-4500, 1988

A:Title: Expression of the *ara* operon in *E. coli* and its role in *ara* biosynthesis

A:Reference number: B31835; MIMD:8908059

A:Accession: B31836

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: EMBL:212147, NID:438134, FIDN:CAAT3131.1, FIL:4167315

Query Match	0.78;	Score 3;	DB 2;	Length 5;
Best Local Similarity	100.0%;	Prod. No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY 352 DTN 454

DB 2 DTN 4

RESULT 21

139964

ribosomal protein S4 - *Bacillus circulans* (fragment)C:Species: *Bacillus circulans*

C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

C:Accession: 139964

R:Grundy, P.J.; Benkin, T.M.

J. Bacteriol. 174, 6763-6770, 1992

A:Title: Characterization of the *Bacillus subtilis* rpsD regulatory target site

A:Reference number: 139964; MIMD:93015735

A:Accession: 139964

A:Status: preliminary; translated from 3c1mp.ppt

A:Molecule type: RNA

A:Residues: 1-5 <RES>

A:Cross-references: DB: M9641; NID:9143471

A:Genes: rpsD

Query Match	0.78;	Score 3;	DB 2;	Length 5;
Best Local Similarity	100.0%;	Prod. No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY 143 MAR 145

DB 1 MAR 3

RESULT 22

139966

ribosomal protein S4 - *Bacillus subtilis* (fragment)C:Species: *Bacillus subtilis*

C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

C:Accession: 139966

R:Grundy, P.J.; Benkin, T.M.

J. Bacteriol. 174, 6763-6770, 1992

A:Title: Characterization of the *Bacillus subtilis* rpsD regulatory target site

A:Reference number: 139964; MIMD:93015735

A:Accession: 139966

A:Status: preliminary; translated from 3c1mp.ppt

A:Molecule type: DNA

A:Residues: 1-5 <RES>

Query Match	0.78;	Score 3;	DB 2;	Length 5;
Best Local Similarity	100.0%;	Prod. No. 2,86-05;		
Matches	3;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY 143 MAR 145

DB 1 MAR 3

Antigen performance (ELISA) - NID-114-175
 (continued)

Antigen type:

Genotype: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 26

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 27

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 28

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 29

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 30

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

Antigen performance (ELISA) - NID-114-175
 (continued)

Antigen type:

Genotype: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 31

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 32

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 33

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 34

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

RESULT 35

100% 100%

Antigen performance (ELISA) - NID-114-175

Antigen type: M27-3

Host: Tissue: Kidney; Tissue: Blood; Nucleic acid: RNA

Method: 1) 100% sensitivity; 2) 100% specificity; 3) 100% accuracy

27 143 MAR 145

100 100

100 100

A:Reference number: S51077, M01D:95045597
 A:Accession: S51077
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 15 CDS

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 OCT 25
 DB 2 OCT 4

RESULT 29
 A44692
 Fulicin - giant African snail
 C:Species: Achatina fulica (giant African snail)
 C>Date: 23-Mar-1997 #sequence_revision 85-Apr-1997 #text_change 11 Jul 1997
 C:Accession: A44692
 R:Ohta, N.; Kihota, I.; Takao, T.; Shimomishi, Y.; Yasuda Kamakura, Y.; Minakata, H.; No
 Biochem. Biophys. Res. Commun. 178, 486-493, 1991
 A>Title: Fulicin, a novel neuropeptide containing a D-amino acid residue isolated from
 A:Reference number: A44692, M01D:91315471
 A:Accession: A44692
 A:Molecule type: protein
 A:Residues: 15 CDS
 C:Keywords: amidated carboxyl end; D-amino acid; neuropeptide
 F12/Modified site: D-aspartoyl (Asp) #status experimental
 F15/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 132 NEF 134
 DB 2 NEF 4

RESULT 30
 150385
 myosin light chain 2 - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 28-Feb-1997
 C:Accession: 150385
 R:Shen, R.; Goswami, S.K.; Mascareno, E.; Kumar, A.; Siddiqui, M.A.Q.
 Mol. Cell. Biol. 11, 1676-1685, 1991
 A>Title: Tissue-specific transcription of the cardiac myosin light-chain 2 gene is regul
 A:Reference number: 150385, M01D:9111519
 A:Accession: 150385
 A:Status: preliminary; translated from GH/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 15 CDS
 A:Cross-references: GB:M63363, MIM:921232
 C:Genetics:
 A:Gene: MYC2

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 165 VSA 167
 DB 2 VSA 4

RESULT 31
 J70520
 19 kappa chain V-III region (SD1) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 16-Aug-1996

C:Accession: J70520
 R:Anker, R.; Conley, M.E.; Pollak, B.A.
 J. Exp. Med. 169, 2109-2119, 1989
 A>Title: Clonal diversity in the B-cell repertoire of patients with X-linked agammag
 A:Reference number: J70520, M01D:96279457
 A:Accession: J70520
 A:Molecule type: mRNA
 A:Residues: 15 CDS
 A>Note: The sequence shown here is one of eight productive V-D-J nu chain rearrange
 A>Note: a stop codon terminates the sequence in the V region
 C:Keywords: heterodimer; immunoglobulin
 F1-5/Domain: V kappa region (VPE)

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 SEP 17
 DB 4 SEP 5

RESULT 32
 S62883
 seminal plasma protein 11 - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
 C:Accession: S62883
 R:Benavente, A.; Varela, P.F.; Sanz, L.; Toepfer-Horsens, F.; Gonzalez, J.L.
 FEBS Lett. 382, 15-17, 1996
 A>Title: 35kDa protein 11 and protein 12 X-ray crystallographic analysis of four seminal p
 A:Reference number: S62883; M01D:96196555
 A:Accession: S62883
 A:Molecule type: protein
 A:Residues: 15 CDS
 C:Complex: heterodimer; seminal plasma protein 1 and seminal plasma protein 11
 C:Keywords: glycoprotein; heterodimer; semen

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 ING 24
 DB 4 ING 5

RESULT 33
 P70610
 T-cell receptor beta chain V-D-J region (100-26) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 20-May-1997
 C:Accession: P70610
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A>Title: Functional analysis of the T cell receptor beta chain V-D-J region from NOD.B6
 A:Reference number: P70610; M01D:91277601
 A:Accession: P70610
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 CDS
 A:Experimental source: newborn thymus strain BALB/c

Query Match 0.78; Score 3; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 SES 170
 DB 4 SES 5

RESULT 64

P10060

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10060

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10060

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 4 postnatal thymus, strain BALB/c

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 54 SSK 65

DB 4 SSK 5

RESULT 65

P10061

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10061

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10061

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 4 postnatal thymus, strain BALB/c

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 166 SSK 166

DB 4 SSK 4

RESULT 66

P10065

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10065

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10065

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: newborn thymus, strain BALB/c

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 176 N60 220

DB 4 N60 5

RESULT 67

P10077

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10077

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10077

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 19 total thymus, strain BALB/c, clone 141 bp

Accession: P10077

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 19 total thymus, strain BALB/c, clone 141 bp

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 166 SSK 166

DB 4 SSK 4

RESULT 68

P10065

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10065

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10065

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 19 total thymus, strain BALB/c

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 166 SSK 166

DB 4 SSK 4

RESULT 69

P10700

1 cell receptor beta chain V beta region (141 bp) mouse (Treatment)

Species: Mus musculus (house mouse)

Cloned 17 Jul 1992 #sequence revision 17 Jul 1992 #next change 30 May 1997

Accession: P10700

Reference: A.L.

1. Exp. Mod. 174, 115-124, 1991

Altitier: functional sequences of total 1 cell receptor beta chains have two N regions

Accession: P10700

Accession: Translation not shown

Accession: Type: mRNA

Accession: 1 % CFE

Experimental source: day 19 total thymus, strain BALB/c

Keywords: 1 cell receptor

Query Match

Best local similarity: 100.0% Ident. No. 2, Score 42

Matches: 42 Contained by: 0 Mismatches: 0 Gaps: 0

QY 166 SSK 166

DB 4 SSK 4

A>Note: lys-1 was also found
C:Keywords: g:yeosyl:transferase, hexosyl:transferase, phosphoprotein
F:4/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 0.78; Score 3; DB 2; Length 6;
Best local Similarity 100.0%; Pref No 2 Rep 0.5;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 ISV 313
|||
DB 3 ISV 5

Search completed: October 9, 2002, 12:44:10
Job time : 26 secs

GenScore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: October 9, 2002, 12:42:21, Search time 12 seconds
(without alignments)
1293.872 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 401
Sequence: 1 MGIQNNVFHCPYISCPYIA.....KKTCNLLVETVYPMWGI 401

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 349

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 1000 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being plotted,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	1.0	7	1	CIA_ENTFA
2	4	1.0	8	1	LCK1_LEDMA
3	4	1.0	8	1	LCK1_LEDMA
4	4	1.0	9	1	COXE_THDOB
5	4	1.0	10	1	NS1_MYCTO
6	3	0.7	3	1	LUXE_VIBF1
7	3	0.7	4	1	CCPI_PCTM1
8	3	0.7	5	1	TPIS_CANFA
9	3	0.7	6	1	ASP2_LACSN
10	3	0.7	6	1	FI01_LITPU
11	3	0.7	6	1	FAP1_MONPY
12	3	0.7	6	1	LCK1_LODMI
13	3	0.7	6	1	QW1_PEPDE
14	3	0.7	7	1	FAP1_ASCSU
15	3	0.7	7	1	FAP4_PANRE
16	3	0.7	8	1	AL16_CARMA
17	3	0.7	8	1	AL16_CARMA
18	3	0.7	8	1	AL11_CVDPO
19	3	0.7	8	1	AL16_CVDPO
20	3	0.7	8	1	AL18_CARMA
21	3	0.7	8	1	AL13_CARMA
22	3	0.7	8	1	R44K_POPGI
23	3	0.7	8	1	FAP4_HOMAM
24	3	0.7	8	1	LCV1_PUMA
25	3	0.7	8	1	LCV1_PUMA
26	3	0.7	8	1	LCV1_PUMA
27	3	0.7	8	1	LCV1_PUMA
28	3	0.7	8	1	LCV1_PUMA
29	3	0.7	8	1	LCV1_PUMA
30	3	0.7	8	1	LCV1_PUMA
31	3	0.7	8	1	LCV1_PUMA
32	3	0.7	8	1	LCV1_PUMA
33	3	0.7	8	1	LCV1_PUMA

[illegible]

[illegible]

325	2	0.5	TEFE	ATHE	12350	antennapedia
327	2	0.5	TEFE	EMEA	12460	radia callosa
329	2	0.5	TEFE	EMEA	12172	pus sericus
330	2	0.5	TEFE	EMEA	12910	phyllocladus
331	2	0.5	TEFE	EMEA	14264	apocis acutif
332	2	0.5	TEFE	EMEA	14265	arabidopsis
333	2	0.5	TEFE	EMEA	14362	urechis and
334	2	0.5	TEFE	EMEA	14318	nicotiana p
335	2	0.5	TEFE	EMEA	14177	leontopodium
336	2	0.5	TEFE	EMEA	14178	leontopodium
337	2	0.5	TEFE	EMEA	14179	leontopodium
338	2	0.5	TEFE	EMEA	14180	leontopodium
339	2	0.5	TEFE	EMEA	14181	leontopodium
340	2	0.5	TEFE	EMEA	14182	leontopodium
341	2	0.5	TEFE	EMEA	14183	leontopodium
342	2	0.5	TEFE	EMEA	14184	leontopodium
343	2	0.5	TEFE	EMEA	14185	leontopodium
344	2	0.5	TEFE	EMEA	14186	leontopodium
345	2	0.5	TEFE	EMEA	14187	leontopodium
346	2	0.5	TEFE	EMEA	14188	leontopodium
347	2	0.5	TEFE	EMEA	14189	leontopodium
348	2	0.5	TEFE	EMEA	14190	leontopodium
349	2	0.5	TEFE	EMEA	14191	leontopodium

```

RESULT 1
C1A_LEMMA          STANDARD      PRI:       7 AA.
ID C1A_LEMMA          P11942;
AC 01-06T-1989 (Rel. 12, Created)
DE 01-06T-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DR Sex pheromone CAM373 (Clumping inducible agent) (CIA).
OS Enterobacteriaceae (Streptococcus faecalis).
OT Bacteria, Firmicutes, Bacillales, Streptococcaceae, Streptococcus,
GX NCBI_FtaxID=1451;
RN [1]
RP SOURCE
EX MEDLINE REFERENCE, PMID: 609276;
FA KATIM, T. & H. E. FROST, Y. IZUMI, A. NISHIOKA, M. KITADA G.,
RT White B.A., An E.Y., Clewell D.B., Suzuki A.T.
RT Association and structure of the Shiga-toxin lactation sex pheromone,
RT CAM373.*?
RE FEBS Lett. 265:69-72(1990);
GZ JOURNALISM, CAM373, INSECT MATING BEHAVIOR 1 FOR K. THIS
CO BARBERINS, FR3373.
GC 1 RECALLS:05; THE N THIRDS IS FREQUENTLY KEPT ASIDE FOR
GZ SPECIFICITY OF PHEROMONES TO PLASMIOS.
GC 1 SIMILARITY, C TERMINAL TO THE PHEROMONE PTD AND PAUL .
DR PIR: A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA: 724 MR: 75EPD20590705UO 70664;

Query Match:              100% Score 4; DR 1; Length 7;
Best Local Similarity:    100.0%; Pred. No. Pos:0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; gaps 0
OY 45 AIFI 4B
DB 1 AIFI 4

```



```

RESULT 6
LUXE_VIBF1
ID LUXE_VIBF1 STANDARD: PRI: 3 AA.
AC P24272
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Long-chain fatty-acyl-luciferin-component ligase (EC 6.2.1.19) (Acyl-
protein synthetase) (fragment).
LUXE:
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9107225; PubMed 2251256;
RA Swartzman E., Kapoor S., Graham A.F., Melnick E.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination
site for the lux operon."
RL Bacteriol 172:6797-6803(1999)
CC -1- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC IT IS A COMPONENT OF THE FATTY ACID ELECTRIC COMPLEX RECOMBINING
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein -> AMP + diphosphate +
CC an acyl-protein thioester.
CC -1- PATHWAY: SECOND STEP IN THE BIOLUMINESCENT FATTY ACID REDUCTASE
CC COMPLEX
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by a third party without
CC explicit permission is prohibited (see http://www.isb-sib.ch/permissions
CC or send an email to license@sib-sib.ch).
CC
CC FMT: M69812; 1 REF: ARN07ACTHT_CDS.
CC
CC Luminescence; Ligase;
CC
CC FT NON_TER
CC
CC SO SEQUENCE 3 AA: 374 MW: 6AA3303000000000 C67644;
Query Match 0.78; Score 3; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 150 kpn 152
DB 1 IKD 3

```

```

RESULT 7
OCP1_OCTM1
ID OCP1_OCTM1 STANDARD: PRI: 4 AA.
AC P58648;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cardioactive peptides OCP-1, OCP-2.
OS Octopus minor (octopus).
OC Eukaryota; Metazoa; Cephalopoda; Coleoida; octopoda;
OC Invertebrata; Cephalopoda; octopus.
OX NCBI_TaxID=89766;
RN [1]
RP SEQUENCE SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX PubMed=10876044;
RA Takashi F., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor."
RL Peptides 21:623-630(2000)
CC -1- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. OCP-2 is a 1000 time less

```

```

CC active than OCP-1.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PIM: OCP-2 has 1 phe instead of 0 phe.
CC -1- MASS SPECTROMETRY: MW 175.2, M3H33 MALDI.
CC Homology: D- amino acid.
CC MOD_RES 2 2 2
CC SEQUENCE 4 AA: 594 MW: 6AAB9C8100000000 C67644;
Query Match 0.78; Score 3; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 FGD 146
DB 2 FGD 4

```

```

RESULT 8
TPIS_CANVA
ID TPIS_CANVA STANDARD: PRI: 6 AA.
AC P54714;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (fragment).
CN (1).
OS Canis familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa;
OC Mammalia; Canidae; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID 9615;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE: 9817340; PubMed=964912;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins."
RL Electrophoresis 19:2725-2823(1998)
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + glycerone
CC phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: RELATES TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
CC HSC-2DPAGE: P54714; ILOC.
DR InterPro: IPR000652; Trioseph_Isomerase.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Glucosephosphate; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER
FT NON_TER
FT NON_TER
SO SEQUENCE 5 AA: 550 MW: 6A4448629A000000 C67644;
Query Match 0.78; Score 3; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GNN 6
DB 2 GNN 5

```

```

RESULT 9
ASP2_LACSN
ID ASP2_LACSN STANDARD: PRI: 6 AA.
AC P91655;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acid shock protein 2 (fragment).
OS Lactobacillus santranciscensis (Lactobacillus santranciscensis).
CC bacteria; Firmicutes; Bacillus; Clostridia; Bacillus; Clostridiaceae;
CC Lactobacillus.
OX NCBI_TaxID 1625;

```


OC Eubrachyura; Portunidae; Portunidae; Carcinus
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion and Thoracic ganglion;
 RX MEDLINE=98121193 PubMed 946295;
 RA Dave H., Johnson A.H., Mastroianni L., Scott A.G., Jares P.D.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RI Eur. J. Biochem. 250:727-734(1997)
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation; Multigene family.
 KW MOD_RES 8
 FT MOD_RES 8
 SO SEQUENCE 8 AA; 781 MW; 7c2879c0c8476878 CIRC64;
 Query Match 0.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 16 GPT 18
 DB 2 GPT 4

RESULT 22

B44K_PORGI
 ID B44K_PORGI STANDARD: PRT; 8 AA.
 AC P81886;
 DT 30-MAY-2000 (Rel. 39, Created)
 RT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 44 KDa Immunogenic protein (Fragment).
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteri; CFU group; Bacteroidetes; Bacteriobas; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=VPB 3492;
 RX MEDLINE=20194497 PubMed=10731616;
 RA Norris J.M., Love D.N.;
 RT "Serum antibody responses of rats to soluble whole cell antigens of
 RT feline Porphyromonas gingivalis.";
 RI Vet. Microbiol. 73:37-49(2000).
 CC -1- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A
 CC Antigen.
 KW NON-TER 8
 FT NON-TER 8
 SO SEQUENCE 8 AA; 984 MW; 95f4540426c9474e CIRC64;

Query Match 0.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e-05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 RNI 298
 DB 6 RNI 8

RESULT 23

FAR4_HOMAM
 ID FAR4_HOMAM STANDARD: PRT; 8 AA.
 AC P41487;
 DT 01-NOV-1995 (Rel. 32, Created)
 RT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRAamide-like neuropeptide 4 (Fli 4) (F1).
 OS Homarus americanus (American lobster).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Decapoda; Decapoda; Fidecyemata; Ascaidea;
 OC Neptopoda; Reptropheta; Homarus
 OX NCBI_TaxID=6706;

RN [1]
 RP SEQUENCE.
 RC TISSUE=Pericardial organs;
 RX MEDLINE=8811644 PubMed 429714;
 RA Timmer E.A., Kobierski L.A., Fliedtz F.A.;
 RT "Pericardial and thoracic isolation of peptide-like immunoreactive
 RT substances from the lobster nervous system: isolation and sequence
 RT analysis of two closely related peptides.";
 RI J. Comp. Neurol. 266:16-26(1987).
 CC -1- FUNCTION: CAN ACT AS A MODULATOR OF EXCITATORY AND CARDIAC
 CC NEUROSMUSCULAR FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE FMRA (FMRA-RELATED PEPTIDE)
 CC FAMILY.
 CC Neuropeptide; Amidation.
 KW MOD_RES 8
 FT MOD_RES 8
 SO SEQUENCE 8 AA; 1067 MW; DDD40729C4540451 CIRC64;
 Query Match 0.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 217 TNR 219
 DB 1 TNR 3

RESULT 24

LCK2_LEUMA
 ID LCK2_LEUMA STANDARD: PRT; 8 AA.
 AC P21141;
 DT 01-MAY-1991 (Rel. 18, Created)
 RT 01-MAY-1991 (Rel. 18, Last sequence update)
 DE Leucokinin II (L-11).
 OS Leucophaea maderae (Mollusca: caudofoveate).
 OC Eutherozoa; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OC Pectiniformia; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OC Bivalvia; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Head;
 RA Holman G.M., Cove R., Neuchman P.J.;
 RT "Isolation, primary structure and synthesis of two neuropeptides
 RT from Leucophaea maderae: members of a new family of
 RT cephalomyotropic peptides.";
 RI Comp. Biochem. Physiol. 104C:205-211(1986).
 CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC ACTIVITY OF COCKROACH PROTHORACIC (HINDGUT).
 CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
 KW Neuropeptide; Amidation.
 KW MOD_RES 8
 FT MOD_RES 8
 SO SEQUENCE 8 AA; 852 MW; D26365A59C6676A CIRC64;

Query Match 0.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e-05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 464 SWG 365
 DB 6 SWG 8

RESULT 25

LCK4_LEUMA
 ID LCK4_LEUMA STANDARD: PRT; 8 AA.
 AC P21143;
 DT 01-MAY-1991 (Rel. 18, Created)
 RT 01-MAY-1991 (Rel. 18, Last sequence update)
 DE Leucokinin IV (L-IV).
 OS Leucophaea maderae (Mollusca: caudofoveate).
 OC Eutherozoa; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OC Pectiniformia; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OC Bivalvia; Mollusca; Arthropoda; Mollusca; Bivalvia; Pectiniformia;
 OX NCBI_TaxID=6988;

AC p199907;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE leucokinin VIII (L-VIII).
 OS leucophaea materae (Mediterranean Cuckoo)
 OC Eukaryota; Metazoa; Arthropoda; Insecta;
 OC Pterygota; Neoptera; Orthoptera; Dictyoptera; Blattaria;
 OC Blaberidae; Blaberidae; Leucophaea.
 OX NCBI_Taxid:6989;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Head;
 RA Holman G.M., Cook R.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of leucokinin VII and
 RT VIII, the final members of this new family of cephalomyotropic
 RT peptides isolated from head extracts of Leucophaea materae.";
 RL Comp Biochem Physiol. 88C:31-34(1987)
 CC -1. FUNCTION THIS CATHAL-MYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 CC ACTIVITY OF CUCKOOB CROCODILUS (HINPRT)
 CC -1. SIMILARITY: TO THE OTHER LEUCOKININS.
 CC PTP: JSD019; JSD019
 DR NeuroPeptide; Amidation
 FT MOD_RES 8
 FT AMINATION.
 SO SEQUENCE 8 AA; 902 MW; 73636FA9C6A9A9A9 CDD64;

Query Match 0.78; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Prod No 1a-0c;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 30			
NPB_BOVIN	STANDARD:	PRTE:	8 AA.
AC P15507;			
DT 01-APR-1990 (rel. 14, Created)			
DT 01-APR-1990 (rel. 14, Last sequence update)			
DT 01-APR-1990 (rel. 14, Last annotation update)			
DR Neuropeptide B.			
OS Bos taurus (Bovine).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Cetartiodactyla; Artiodactyla; Perissodactyla; Bovidae;			
OC Bovidae; Bovinae; Bos.			
OX NCBL_TaxID:9913;			
LN [1]			
RP SEQUENCE:			
RP TISSUE=Brain;			
RX MEDLINE=6606/985; PubMed=3665193;			
RA Yang H.-Y.T., Fratta W., Majum F.A., Costa E.;			
RT "Isolation, sequencing, synthesis, and pharmacological			
RT characterization of two brain neuropeptides that modulate the action			
RT of morphine.";			
RT			
PI Proc. Natl. Acad. Sci. U S A. 82:7767-7761(1985).			
CC - FUNCTION: MODULATES THE ACTION OF MORPHINE.			
CC PIP: R24749; R24749.			
CC Neuropeptide; Amidation.			
ET SIMILAR 5 8 TO NEUROPEPTIDE A (AA 5-8) (IDENTICAL).			
ET MOD_RES 8 8 AMIDATION.			
SEQUENCE 4 AA: 1082 MW: 6714.6677660726 (79064;			

RESULT	31
ORMY_ORCL1	
ID	ORMY_ORCL1
AC	P82455;
DT	16-OCT-2001 (Ref. 40, Created)
ET	15-OCT-2001 (Ref. 40, Last sequence update)
DT	16-OCT-2001 (Ref. 40, Last annotation update)
DE	Oreomyzotropin (IMT).
OS	Oreomyzetes limosus (Spinthaeck crassish).
OC	Eukaryota, Metazoa, Arthropoda, Crustacea, Malacostraca;
OC	Eumalacostraca; Eucarida; Decapoda, Pleocyemata; Astacidea;
OC	Astacoidae; Cambaridae; Oreomyzetes.
OX	NCBI_ProtID=28379;
RN	[1]
RP	SEQUENCE, MASS SPECTROMETRY, AND AMIDATION.
RC	1153bp-Hindcut;
FX	MEDLINE=20411510; PubMed=10952880;
PA	Direksen H., Rørdal S., Sænter A., Keller R.;
FT	"Two oreklimins and two novel oreklimin-like myotropin in the hindgut
FT	of the crayfish oreomyzetes limosus: identification of myotropin
FT	neuropeptides originating together in neurons of the terminal
FT	abdominal ganglion.";
JL	J. Exp. Biol. 203:2267-2318(2000).
CC	-1- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
CC	AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS. IT IS SYNTHESIZED
CC	BY ABDOMINAL GANGLIONIC NEURONS.
CC	-1- MASS SPECTROMETRY: MM-904.8. METHOD-FAH.
FW	Amidation; Neuropeptide.
FT	MOD_RES 8 AA: 8
FT	MOD_RES 8 AA: 905 MW; RT(0.018)A(0.000)A(0.004);

```

RESULT 32
RS1_ERWCH          STANDARD          PR1          8 AA.
ID_RS1_ERWCH          P37985
AC          01-OCT-1994 (Rel. 30, Created)
DT          01-OCT-1994 (Rel. 30, Last sequence update)
DE          01-NOV-1995 (Rel. 32, Last annotation update)
DR          30S ribosomal protein S1 (Fragment) .
GN          RPSA.
OS          Erwinia chrysanthemi.
OC          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae
OC          Pectobacterium.
OX          NCBI_TaxID:556;
[1]
SEQUENCE FROM N.A.
RP          STRAIN 4937:
RC          Doullie A., Toussein A., Paalen M.;
RA          J. Antibiot. (1993) 46, 1063-1065.
CC          ! PROTEIN, ERWCH ERWA, ERWG FACILITATE PRESENTATION OF THE
CC          ! INITIATION POINT. IT IS NECESSARY TO TRANSLATE ERWA WITH A SH R
CC          SHIM (ERL280) (SI) FOR THE RICH SEQUENCE (BY SIMILARITY)
CC          -1- SIMILARITY BELONGS TO THE S16 FAMILY OF RIBOSOMAL PROTEINS.
CC          -----
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL institution.
CC          The European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed, usage by and for commercial
CC          entities requires a license agreement with the EMBL institution.
CC          or send an email to license@isb.slb.ch).
CC          -----
CC          EMBL: X74750; CAA52769.1;
CC          PIR: S37141; S37141.
DR

```


DT 01-NOV-1998 (Rel. 09, Created)
 DT 01-NOV-1998 (Rel. 09, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Lys-conopressin G.
 OS Conus geographus (Geography Conus).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus
 NCBI_TaxID=6491;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=88058932; PubMed=3680228;
 RA Cruz L.J., de Santos V., Zataralla G.C., Ramirez C.A., Zelus R.D.,
 RA Gray W.R., Olivera B.M.;
 RT "Invertebrate vasopressin/oxytocin homologs: characterization of
 RT peptides from Conus geographus and Conus striatus venoms.";
 RL J. Biol. Chem. 262:15821-15824(1987)
 RN [2]
 RP REVIEW.
 RX MEDLINE=89034584; PubMed=3692984;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988)
 CC -1- SIMILARITY: RELIGIONS TO THE VASOPRESSIN/OXYTOCIN FAMILY
 DR PIR: A28495; A28495;
 DE InterPro: IPR000981; Neurohypophys_horm.
 DR Pfam: PF00220; hormone4; 1
 DR PROSITE: PS00264; Neurohypophys_horm; 1
 KW Hormone; Amidation.
 FT DISULFID 1 6
 FT MOD_RES 9 9
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA: 1037 MW: 144927684.4059 cpk64.

Query Match 0.78; Score 37, PP 13, Length 9;
 Best local Similarity 100.0%; Prod. No. 1e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 RN 5

RESULT 37
 ID CONO_CONST STANDARD; PRT; 9 AA.
 AC POS487;
 DT 01-NOV-1998 (Rel. 09, Created)
 DT 01-NOV-1998 (Rel. 09, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Arg-conopressin S.
 OS Conus striatus (Striated Conus).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus
 NCBI_TaxID=6493;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=88058932; PubMed=3680228;
 RA Cruz L.J., de Santos V., Zataralla G.C., Ramirez C.A., Zelus R.D.,
 RA Gray W.R., Olivera B.M.;
 RT "Invertebrate vasopressin/oxytocin homologs: characterization of
 RT peptides from Conus geographus and Conus striatus venoms.";
 RL J. Biol. Chem. 262:15821-15824(1987).
 RN [2]
 RP REVIEW.
 RX MEDLINE=89034584; PubMed=3692984;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988)
 CC -1- SIMILARITY: RELIGIONS TO THE VASOPRESSIN/OXYTOCIN FAMILY
 DR PIR: B28495; B28495;
 DE InterPro: IPR000981; Neurohypophys_horm
 DR Pfam: PF00220; hormone4; 1
 DR PROSITE: PS00264; Neurohypophys_horm; 1
 KW Hormone; Amidation.

FT DISULFID 1 6
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA: 1031 MW: 173817684.540050 cpk64;
 Query Match 0.78; Score 37, PP 13, Length 9;
 Best local Similarity 100.0%; Prod. No. 1e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 RN 5

RESULT 38
 ID DNF1_LINKM STANDARD; PRT; 9 AA.
 AC P16339;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Locopressin (diuretic neuropeptide F1/72).
 OS Locusta migratoria (migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 OC Orthoptera; Orthoptera; Orthoptera; Orthoptera; Orthoptera;
 OC Acridomorpha; Acrididae; Acrididae; Locusta.
 NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RX TISSUE=Suboesophageal ganglion; and Thoracic ganglion;
 RX MEDLINE=88077077; PubMed=4684410;
 RA Pavesi F., Mitter C.A., Mitter C.A., Mitter C.A., Mitter C.A.,
 RA Delage M., Schooley D.A.;
 RT "Identification of an inhibitory vasopressin-like diuretic hormone from
 RT Locusta migratoria.";
 RL Biochem. Biophys. Res. Commun. 114:180-186(1983).
 CC -1- FUNCTION: DIURETIC HORMONE.
 CC -1- FUNCTION: PZ IS AN ANTIHYPERTENSIVE DISULFIDE LINKED MEMBER OF FL.
 CC -1- SIMILARITY: RELIGIONS TO THE VASOPRESSIN/OXYTOCIN FAMILY
 DR PIR: A29477; A29477;
 DE InterPro: IPR000981; Neurohypophys_horm.
 DR Pfam: PF00220; hormone4; 1
 DR PROSITE: PS00264; Neurohypophys_horm; 1
 KW Hormone; Neurohypophyside; Amidation.
 FT DISULFID 1 6
 FT DISULFID 1 1
 FT MOD_RES 9 9
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA: 976 MW: 549372584.1457 cpk64;

Query Match 0.78; Score 37, PP 13, Length 9;
 Best local Similarity 100.0%; Prod. No. 1e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 RN 5

RESULT 39
 ID FAR5_ASCSU STANDARD; PRT; 9 AA.
 AC P41170;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE FMRFamide-like neuropeptide AF5.
 OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridida;
 OC Ascarididae; Ascaris.
 NCBI_TaxID=6253;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95380742; PubMed=7651904;

ID OXYT_EISFO STANDARD: PRT: 9 AA.
AC P42998;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 01-NOV-1995 (Rel. 32, last annotation update)
DE Annelocin.
OS Eisenia foetida (Common branding worm) (Common dung-worm);
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6366;
RN [1]
RP SEQUENCE:
RC TISSUE=Plutary;
RX MEDLINE=94121660; PubMed 8292046;
RA Oumi T., Ukeda K., Matsushima O., Ikeda T., Fujita T., Minakata H.,
RA Nomoto K.;
RT "Annelocin: an oxytocin-related peptide isolated from the earthworm,
RT Eisenia foetida.";
RL Biochem. Biophys. Res. Commun. 198; 393-396(1994).
CC -1- FUNCTION: POTENTIATES SPONTANEOUS CONTRACTIONS OF THE GUT AND ALSO
CC PUSATORY CONTRACTIONS AND BLADDER-SHAKING MOVEMENT OF THE
CC NEPHRIDIAL FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR PUBMED: P002021; P002021.
DR InterPro: IPR000981; Neutrophys_horm.
DR Pfam: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; FALSE_NEG.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
FT SEQUENCE 9 AA: 996 MW: 146876684541209 CRC64;

Query Match 0.78; Score 3; DB 1; length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 360 VRN 362
|||
Db 3 VRN 5

RESULT 50
OXYT_RAJCL STANDARD: PRT: 9 AA.
ID OXYT_RAJCL
AC P42994;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 01-NOV-1995 (Rel. 32, last annotation update)
DE Glutocin.
OS Raja clavata (Thornback ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Hypnosqualea; Plistiorajea; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7781;
RN [1]
RP SEQUENCE:
RX MEDLINE=66123415; PubMed 5880565;
RA Acher R., Chauvet J., Chauvet M.-T., Creppy D.;
RT "Physiology of neurohypophyseal peptides: isolation of a new hormone,
RT glutocin (ser 4-61n 8-oxytocin) present in a cartilaginous fish,
RT the ray (Raja clavata).";
RL Biochim. Biophys. Acta 107; 393-396(1985).
CC -1- FUNCTION: ANTIHYPERTENSIVE HORMONE.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR InterPro: IPR000981; Neutrophys_horm.
DR Pfam: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
FT SEQUENCE 9 AA: 984 MW: 179976684558048 CRC64;

Query Match 0.78; Score 3; DB 1; length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 YIS 15
|||
Db 2 YIS 4

Search completed: October 9, 2002, 12:42:50
Job time : 14 secs



GenCorp version 5.1.3
Copyright (c) 1993 - 2002 Computer Ltd.

OM protein : protein search, using sw model

Run on: October 9, 2002, 12:42:22 / Search time 29 Seconds

(without alignments)
2392.102 Million cell updates/sec

Title: US-09-598-062-4

Percent score: 401

Sequence: 1 MDIGNVHERHGEISGPVIA

KKTCGIIIVPYVIMNSI 401

Scoring table:

Gapop 60.0, Gapext 60.0

Searched:

552222 seqs, 12294129 residues

Word size:

0

Total number of hits satisfying chosen parameters:

1088

Minimum DB seq length: 0

Maximum DB seq length: 10

Post processing: Listing first 1000 summaries

Database:

1: SPREMI_19: *
2: sp.archaea: *
3: sp.bacteria: *
4: sp.fungi: *
5: sp.human: *
6: sp.invertebrate: *
7: sp.mammal: *
8: sp.mbc: *
9: sp.organeller: *
10: sp.plant: *
11: sp.podent: *
12: sp.virus: *
13: sp.vertebrate: *
14: sp.unclassified: *
15: sp.virus: *
16: sp.bacteriaph: *
17: sp.archaea: *

Prod No is the number of results predicted by choice to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	1.0	7	11	Q63668
2	4	1.0	8	2	Q63668
3	4	1.0	8	2	Q63668
4	4	1.0	8	5	Q96069
5	4	1.0	8	5	Q96065
6	4	1.0	8	5	P82685
7	4	1.0	8	5	P82687
8	4	1.0	8	5	P82689
9	4	1.0	8	8	Q94008
10	4	1.0	8	9	Q37854
11	4	1.0	8	11	P70243
12	4	1.0	9	4	Q90634
13	4	1.0	9	10	Q94XHR
14	4	1.0	9	12	Q65545
15	4	1.0	10	10	P82938
16	4	1.0	10	10	Q94787

17	4	1.0	10	11	Q96VJ6	Q96VJ6 rat
18	4	1.0	10	11	Q96VJ6	Q96VJ6 rat
19	4	0.7	6	13	P82096	P82096 human
20	3	0.7	7	2	Q07354	Q07354 human
21	4	0.7	7	2	Q47029	Q47029 human
22	4	0.7	7	2	Q59556	Q59556 human
23	4	0.7	7	2	P70804	P70804 human
24	4	0.7	7	2	Q51219	Q51219 human
25	4	0.7	7	10	P93423	P93423 human
26	4	0.7	7	11	P92445	P92445 human
27	4	0.7	7	11	Q63480	Q63480 human
28	4	0.7	8	2	Q66580	Q66580 human
29	4	0.7	8	2	Q66580	Q66580 human
30	4	0.7	8	2	Q66517	Q66517 human
31	4	0.7	8	2	Q932K1	Q932K1 human
32	4	0.7	8	2	Q87471	Q87471 human
33	4	0.7	8	2	Q92E79	Q92E79 human
34	4	0.7	8	2	Q53790	Q53790 human
35	4	0.7	8	2	Q49511	Q49511 human
36	4	0.7	8	2	Q56429	Q56429 human
37	4	0.7	8	2	Q56140	Q56140 human
38	4	0.7	8	2	Q56759	Q56759 human
39	4	0.7	8	2	Q62484	Q62484 human
40	4	0.7	8	3	Q90R89	Q90R89 human
41	4	0.7	8	3	Q9P285	Q9P285 human
42	4	0.7	8	4	Q9R600	Q9R600 human
43	4	0.7	8	4	Q9P0K3	Q9P0K3 human
44	4	0.7	8	4	Q9BY95	Q9BY95 human
45	4	0.7	8	4	Q15898	Q15898 human
46	4	0.7	8	4	Q15902	Q15902 human
47	4	0.7	8	4	Q66965	Q66965 human
48	4	0.7	8	5	Q15896	Q15896 human
49	4	0.7	8	5	P82686	P82686 human
50	4	0.7	8	5	P82688	P82688 human
51	4	0.7	8	6	Q91R88	Q91R88 human
52	4	0.7	8	6	Q18854	Q18854 human
53	4	0.7	8	6	Q91178	Q91178 human
54	4	0.7	8	6	Q9HFA7	Q9HFA7 human
55	4	0.7	8	6	Q9HFA0	Q9HFA0 human
56	4	0.7	8	6	Q9HFA9	Q9HFA9 human
57	4	0.7	8	6	P82929	P82929 human
58	4	0.7	8	6	Q95M23	Q95M23 human
59	4	0.7	8	8	Q91280	Q91280 human
60	4	0.7	8	8	Q91002	Q91002 human
61	4	0.7	8	8	Q9MD43	Q9MD43 human
62	4	0.7	8	8	Q94V53	Q94V53 human
63	4	0.7	8	8	Q94V60	Q94V60 human
64	4	0.7	8	8	Q94PX7	Q94PX7 human
65	4	0.7	8	9	Q51R93	Q51R93 human
66	4	0.7	8	6	Q64835	Q64835 human
67	4	0.7	8	10	Q58337	Q58337 human
68	4	0.7	8	10	Q9S824	Q9S824 human
69	4	0.7	8	10	Q40540	Q40540 human
70	4	0.7	8	10	P82324	P82324 human
71	4	0.7	8	11	Q96V15	Q96V15 human
72	4	0.7	8	11	Q91107	Q91107 human
73	4	0.7	8	11	Q99P40	Q99P40 human
74	4	0.7	8	11	Q62528	Q62528 human
75	4	0.7	8	11	P82598	P82598 human
76	4	0.7	8	12	Q84271	Q84271 human
77	4	0.7	8	12	Q83927	Q83927 human
78	4	0.7	8	12	Q83332	Q83332 human
79	4	0.7	8	12	Q9M133	Q9M133 human
80	4	0.7	8	12	Q64971	Q64971 human
81	4	0.7	8	12	Q91021	Q91021 human
82	4	0.7	8	12	Q91019	Q91019 human
83	4	0.7	8	13	P87488	P87488 human
84	4	0.7	8	13	Q90498	Q90498 human
85	4	0.7	8	13	Q91098	Q91098 human
86	4	0.7	8	13	Q92975	Q92975 human
87	4	0.7	8	15	Q85562	Q85562 human
88	4	0.7	8	1	Q60832	Q60832 human
89	4	0.7	9	2	Q96635	Q96635 human

236	3	0.7	10	8	Q9T8S7	Q9T8S7	1101acmus	h	309	4	0.7	10	12	Q9G0N5	Q9G0N5	polyomavirus
237	3	0.7	10	8	Q9T8S4	Q9T8S4	1101acmus	c	310	4	0.7	10	12	Q9G0M3	Q9G0M3	polyomavirus
238	3	0.7	10	8	Q9T8S1	Q9T8S1	1101acmus	1	311	4	0.7	10	12	Q9G0M1	Q9G0M1	polyomavirus
239	3	0.7	10	8	Q9T8P9	Q9T8P9	1101acmus	b	312	4	0.7	10	12	Q9G0V9	Q9G0V9	polyomavirus
240	3	0.7	10	8	Q9T8P7	Q9T8P7	1101acmus	c	313	4	0.7	10	12	Q9G0V7	Q9G0V7	polyomavirus
241	3	0.7	10	8	Q9T8R4	Q9T8R4	1101acmus	p	314	4	0.7	10	13	Q9G0Y8	Q9G0Y8	polyomavirus
242	3	0.7	10	8	Q9T8P1	Q9T8P1	1101acmus	1	315	4	0.7	10	13	Q9G0Y9	Q9G0Y9	polyomavirus
243	3	0.7	10	8	Q9T8O8	Q9T8O8	1101acmus	c	316	4	0.7	10	13	Q9G0U1	Q9G0U1	polyomavirus
244	3	0.7	10	8	Q9T8O5	Q9T8O5	1101acmus	1	317	4	0.7	10	13	Q9T8W9	Q9T8W9	polyomavirus
245	3	0.7	10	8	Q9T8Q2	Q9T8Q2	1101acmus	3	318	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
246	3	0.7	10	8	Q9T8P9	Q9T8P9	1101acmus	3	319	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
247	3	0.7	10	8	Q9T8P6	Q9T8P6	1101acmus	3	320	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
248	3	0.7	10	8	Q9T8P3	Q9T8P3	1101acmus	3	321	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
249	3	0.7	10	8	Q9T8P0	Q9T8P0	1101acmus	3	322	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
250	3	0.7	10	8	Q9T8M7	Q9T8M7	1101acmus	3	323	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
251	3	0.7	10	8	Q9T8M4	Q9T8M4	1101acmus	d	324	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
252	3	0.7	10	8	Q9T8M1	Q9T8M1	1101acmus	f	325	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
253	3	0.7	10	8	Q9T8M8	Q9T8M8	1101acmus	m	326	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
254	3	0.7	10	8	Q9T8M5	Q9T8M5	1101acmus	3	327	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
255	3	0.7	10	8	Q9T8M2	Q9T8M2	1101acmus	3	328	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
256	3	0.7	10	8	Q9T8L9	Q9T8L9	1101acmus	3	329	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
257	3	0.7	10	8	Q9T8L6	Q9T8L6	1101acmus	3	330	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
258	3	0.7	10	8	Q9T8L3	Q9T8L3	1101acmus	3	331	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
259	3	0.7	10	8	Q9T8L0	Q9T8L0	1101acmus	3	332	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
260	3	0.7	10	8	Q9T8K7	Q9T8K7	1101acmus	3	333	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
261	3	0.7	10	8	Q9T8K4	Q9T8K4	1101acmus	3	334	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
262	3	0.7	10	8	Q9T8K1	Q9T8K1	1101acmus	3	335	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
263	3	0.7	10	8	Q9T8J8	Q9T8J8	1101acmus	3	336	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
264	3	0.7	10	8	Q9T8J5	Q9T8J5	1101acmus	3	337	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
265	3	0.7	10	8	Q9T8J2	Q9T8J2	1101acmus	3	338	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
266	3	0.7	10	8	Q9T8J9	Q9T8J9	1101acmus	3	339	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
267	3	0.7	10	8	Q9T8J6	Q9T8J6	1101acmus	3	340	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
268	3	0.7	10	8	Q9T8J3	Q9T8J3	1101acmus	3	341	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
269	3	0.7	10	8	Q9T8J0	Q9T8J0	1101acmus	3	342	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
270	3	0.7	10	8	Q9T8H7	Q9T8H7	1101acmus	3	343	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
271	3	0.7	10	8	Q9T8H4	Q9T8H4	1101acmus	3	344	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
272	3	0.7	10	8	Q9T8H1	Q9T8H1	1101acmus	3	345	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
273	3	0.7	10	8	Q9T8G8	Q9T8G8	1101acmus	3	346	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
274	3	0.7	10	8	Q9T8G5	Q9T8G5	1101acmus	3	347	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
275	3	0.7	10	8	Q9T8G2	Q9T8G2	1101acmus	3	348	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
276	3	0.7	10	8	Q9T8G9	Q9T8G9	1101acmus	3	349	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
277	3	0.7	10	8	Q9T8G6	Q9T8G6	1101acmus	3	350	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
278	3	0.7	10	8	Q9T8G3	Q9T8G3	1101acmus	3	351	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
279	3	0.7	10	8	Q9T8G0	Q9T8G0	1101acmus	3	352	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
280	3	0.7	10	8	Q9T8F7	Q9T8F7	1101acmus	3	353	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
281	3	0.7	10	8	Q9T8F4	Q9T8F4	1101acmus	3	354	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
282	3	0.7	10	8	Q9T8F1	Q9T8F1	1101acmus	3	355	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
283	3	0.7	10	8	Q9T8E8	Q9T8E8	1101acmus	3	356	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
284	3	0.7	10	8	Q9T8E5	Q9T8E5	1101acmus	3	357	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
285	3	0.7	10	8	Q9T8E2	Q9T8E2	1101acmus	3	358	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
286	3	0.7	10	8	Q9T8E9	Q9T8E9	1101acmus	3	359	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
287	3	0.7	10	8	Q9T8E6	Q9T8E6	1101acmus	3	360	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
288	3	0.7	10	8	Q9T8E3	Q9T8E3	1101acmus	3	361	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
289	3	0.7	10	8	Q9T8E0	Q9T8E0	1101acmus	3	362	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
290	3	0.7	10	8	Q9T8D7	Q9T8D7	1101acmus	3	363	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
291	3	0.7	10	8	Q9T8D4	Q9T8D4	1101acmus	3	364	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
292	3	0.7	10	8	Q9T8D1	Q9T8D1	1101acmus	3	365	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
293	3	0.7	10	8	Q9T8C8	Q9T8C8	1101acmus	3	366	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
294	3	0.7	10	8	Q9T8C5	Q9T8C5	1101acmus	3	367	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
295	3	0.7	10	8	Q9T8C2	Q9T8C2	1101acmus	3	368	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
296	3	0.7	10	8	Q9T8C9	Q9T8C9	1101acmus	3	369	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
297	3	0.7	10	8	Q9T8C6	Q9T8C6	1101acmus	3	370	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
298	3	0.7	10	8	Q9T8C3	Q9T8C3	1101acmus	3	371	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus
299	3	0.7	10	8	Q9T8C0	Q9T8C0	1101acmus	3	372	4	0.7	10	13	Q9T8S7	Q9T8S7	polyomavirus
300	3	0.7	10	8	Q9T8B7	Q9T8B7	1101acmus	3	373	4	0.7	10	13	Q9T8S8	Q9T8S8	polyomavirus
301	3	0.7	10	8	Q9T8B4	Q9T8B4	1101acmus	3	374	4	0.7	10	13	Q9T8S9	Q9T8S9	polyomavirus
302	3	0.7	10	8	Q9T8B1	Q9T8B1	1101acmus	3	375	4	0.7	10	13	Q9T8S0	Q9T8S0	polyomavirus
303	3	0.7	10	8	Q9T8A8	Q9T8A8	1101acmus	3	376	4	0.7	10	13	Q9T8S1	Q9T8S1	polyomavirus
304	3	0.7	10	8	Q9T8A5	Q9T8A5	1101acmus	3	377	4	0.7	10	13	Q9T8S2	Q9T8S2	polyomavirus
305	3	0.7	10	8	Q9T8A2	Q9T8A2	1101acmus	3	378	4	0.7	10	13	Q9T8S3	Q9T8S3	polyomavirus
306	3	0.7	10	8	Q9T8A9	Q9T8A9	1101acmus	3	379	4	0.7	10	13	Q9T8S4	Q9T8S4	polyomavirus
307	3	0.7	10	8	Q9T8A6	Q9T8A6	1101acmus	3	380	4	0.7	10	13	Q9T8S5	Q9T8S5	polyomavirus
308	3	0.7	10	8	Q9T8A3	Q9T8A3	1101acmus	3	381	4	0.7	10	13	Q9T8S6	Q9T8S6	polyomavirus

528	2	0.5	R	R	p93957	p93957	testis	501	2	0.5	R	12	Q9805	Q9805	testis
529	2	0.5	R	R	Q19956	Q19956	testis	502	2	0.5	R	12	Q9804	Q9804	testis
530	2	0.5	R	R	Q19957	Q19957	testis	503	2	0.5	R	12	Q9803	Q9803	testis
531	2	0.5	R	R	Q19958	Q19958	testis	504	2	0.5	R	12	Q9802	Q9802	testis
532	2	0.5	R	R	Q19959	Q19959	testis	505	2	0.5	R	12	Q9801	Q9801	testis
533	2	0.5	R	R	Q19960	Q19960	testis	506	2	0.5	R	12	Q9800	Q9800	testis
534	2	0.5	R	R	Q19961	Q19961	testis	507	2	0.5	R	12	Q9799	Q9799	testis
535	2	0.5	R	R	Q19962	Q19962	testis	508	2	0.5	R	12	Q9798	Q9798	testis
536	2	0.5	R	R	Q19963	Q19963	testis	509	2	0.5	R	12	Q9797	Q9797	testis
537	2	0.5	R	R	Q19964	Q19964	testis	510	2	0.5	R	12	Q9796	Q9796	testis
538	2	0.5	R	R	Q19965	Q19965	testis	511	2	0.5	R	12	Q9795	Q9795	testis
539	2	0.5	R	R	Q19966	Q19966	testis	512	2	0.5	R	12	Q9794	Q9794	testis
540	2	0.5	R	R	Q19967	Q19967	testis	513	2	0.5	R	12	Q9793	Q9793	testis
541	2	0.5	R	R	Q19968	Q19968	testis	514	2	0.5	R	12	Q9792	Q9792	testis
542	2	0.5	R	R	Q19969	Q19969	testis	515	2	0.5	R	12	Q9791	Q9791	testis
543	2	0.5	R	R	Q19970	Q19970	testis	516	2	0.5	R	12	Q9790	Q9790	testis
544	2	0.5	R	R	Q19971	Q19971	testis	517	2	0.5	R	12	Q9789	Q9789	testis
545	2	0.5	R	R	Q19972	Q19972	testis	518	2	0.5	R	12	Q9788	Q9788	testis
546	2	0.5	R	R	Q19973	Q19973	testis	519	2	0.5	R	12	Q9787	Q9787	testis
547	2	0.5	R	R	Q19974	Q19974	testis	520	2	0.5	R	12	Q9786	Q9786	testis
548	2	0.5	R	R	Q19975	Q19975	testis	521	2	0.5	R	12	Q9785	Q9785	testis
549	2	0.5	R	R	Q19976	Q19976	testis	522	2	0.5	R	12	Q9784	Q9784	testis
550	2	0.5	R	R	Q19977	Q19977	testis	523	2	0.5	R	12	Q9783	Q9783	testis
551	2	0.5	R	R	Q19978	Q19978	testis	524	2	0.5	R	12	Q9782	Q9782	testis
552	2	0.5	R	R	Q19979	Q19979	testis	525	2	0.5	R	12	Q9781	Q9781	testis
553	2	0.5	R	R	Q19980	Q19980	testis	526	2	0.5	R	12	Q9780	Q9780	testis
554	2	0.5	R	R	Q19981	Q19981	testis	527	2	0.5	R	12	Q9779	Q9779	testis
555	2	0.5	R	R	Q19982	Q19982	testis	528	2	0.5	R	12	Q9778	Q9778	testis
556	2	0.5	R	R	Q19983	Q19983	testis	529	2	0.5	R	12	Q9777	Q9777	testis
557	2	0.5	R	R	Q19984	Q19984	testis	530	2	0.5	R	12	Q9776	Q9776	testis
558	2	0.5	R	R	Q19985	Q19985	testis	531	2	0.5	R	12	Q9775	Q9775	testis
559	2	0.5	R	R	Q19986	Q19986	testis	532	2	0.5	R	12	Q9774	Q9774	testis
560	2	0.5	R	R	Q19987	Q19987	testis	533	2	0.5	R	12	Q9773	Q9773	testis
561	2	0.5	R	R	Q19988	Q19988	testis	534	2	0.5	R	12	Q9772	Q9772	testis
562	2	0.5	R	R	Q19989	Q19989	testis	535	2	0.5	R	12	Q9771	Q9771	testis
563	2	0.5	R	R	Q19990	Q19990	testis	536	2	0.5	R	12	Q9770	Q9770	testis
564	2	0.5	R	R	Q19991	Q19991	testis	537	2	0.5	R	12	Q9769	Q9769	testis
565	2	0.5	R	R	Q19992	Q19992	testis	538	2	0.5	R	12	Q9768	Q9768	testis
566	2	0.5	R	R	Q19993	Q19993	testis	539	2	0.5	R	12	Q9767	Q9767	testis
567	2	0.5	R	R	Q19994	Q19994	testis	540	2	0.5	R	12	Q9766	Q9766	testis
568	2	0.5	R	R	Q19995	Q19995	testis	541	2	0.5	R	12	Q9765	Q9765	testis
569	2	0.5	R	R	Q19996	Q19996	testis	542	2	0.5	R	12	Q9764	Q9764	testis
570	2	0.5	R	R	Q19997	Q19997	testis	543	2	0.5	R	12	Q9763	Q9763	testis
571	2	0.5	R	R	Q19998	Q19998	testis	544	2	0.5	R	12	Q9762	Q9762	testis
572	2	0.5	R	R	Q19999	Q19999	testis	545	2	0.5	R	12	Q9761	Q9761	testis
573	2	0.5	R	R	Q20000	Q20000	testis	546	2	0.5	R	12	Q9760	Q9760	testis
574	2	0.5	R	R	Q20001	Q20001	testis	547	2	0.5	R	12	Q9759	Q9759	testis
575	2	0.5	R	R	Q20002	Q20002	testis	548	2	0.5	R	12	Q9758	Q9758	testis
576	2	0.5	R	R	Q20003	Q20003	testis	549	2	0.5	R	12	Q9757	Q9757	testis
577	2	0.5	R	R	Q20004	Q20004	testis	550	2	0.5	R	12	Q9756	Q9756	testis
578	2	0.5	R	R	Q20005	Q20005	testis	551	2	0.5	R	12	Q9755	Q9755	testis
579	2	0.5	R	R	Q20006	Q20006	testis	552	2	0.5	R	12	Q9754	Q9754	testis
580	2	0.5	R	R	Q20007	Q20007	testis	553	2	0.5	R	12	Q9753	Q9753	testis
581	2	0.5	R	R	Q20008	Q20008	testis	554	2	0.5	R	12	Q9752	Q9752	testis
582	2	0.5	R	R	Q20009	Q20009	testis	555	2	0.5	R	12	Q9751	Q9751	testis
583	2	0.5	R	R	Q20010	Q20010	testis	556	2	0.5	R	12	Q9750	Q9750	testis
584	2	0.5	R	R	Q20011	Q20011	testis	557	2	0.5	R	12	Q9749	Q9749	testis
585	2	0.5	R	R	Q20012	Q20012	testis	558	2	0.5	R	12	Q9748	Q9748	testis
586	2	0.5	R	R	Q20013	Q20013	testis	559	2	0.5	R	12	Q9747	Q9747	testis
587	2	0.5	R	R	Q20014	Q20014	testis	560	2	0.5	R	12	Q9746	Q9746	testis
588	2	0.5	R	R	Q20015	Q20015	testis	561	2	0.5	R	12	Q9745	Q9745	testis
589	2	0.5	R	R	Q20016	Q20016	testis	562	2	0.5	R	12	Q9744	Q9744	testis
590	2	0.5	R	R	Q20017	Q20017	testis	563	2	0.5	R	12	Q9743	Q9743	testis
591	2	0.5	R	R	Q20018	Q20018	testis	564	2	0.5	R	12	Q9742	Q9742	testis
592	2	0.5	R	R	Q20019	Q20019	testis	565	2	0.5	R	12	Q9741	Q9741	testis
593	2	0.5	R	R	Q20020	Q20020	testis	566	2	0.5	R	12	Q9740	Q9740	testis
594	2	0.5	R	R	Q20021	Q20021	testis	567	2	0.5	R	12	Q9739	Q9739	testis
595	2	0.5	R	R	Q20022	Q20022	testis	568	2	0.5	R	12	Q9738	Q9738	testis
596	2	0.5	R	R	Q20023	Q20023	testis	569	2	0.5	R	12	Q9737	Q9737	testis
597	2	0.5	R	R	Q20024	Q20024	testis	570	2	0.5	R	12	Q9736	Q9736	testis
598	2	0.5	R	R	Q20025	Q20025	testis	571	2	0.5	R	12	Q9735	Q9735	testis
599	2	0.5	R	R	Q20026	Q20026	testis	572	2	0.5	R	12	Q9734	Q9734	testis
600	2	0.5	R	R	Q20027	Q20027	testis	573	2	0.5	R	12	Q9733	Q9733	testis

111	09/1/2	PRELIMINARY	1612	AA
110	09/1/2			
109	MAY 2000 (11/2000) 14, 17 (10/00)			
108	MAY 2000 (11/2000) 14, 17 (10/00)			
107	09/2000 (11/2000) 19, 19 (10/00)			
106	11/2000 (11/2000) 19, 19 (10/00)			
105	Y-2000			
104	11/2000 (11/2000) 19, 19 (10/00)			
103	11/2000 (11/2000) 19, 19 (10/00)			
102	11/2000 (11/2000) 19, 19 (10/00)			
101	11/2000 (11/2000) 19, 19 (10/00)			
100	11/2000 (11/2000) 19, 19 (10/00)			
99	11/2000 (11/2000) 19, 19 (10/00)			
98	11/2000 (11/2000) 19, 19 (10/00)			
97	11/2000 (11/2000) 19, 19 (10/00)			
96	11/2000 (11/2000) 19, 19 (10/00)			
95	11/2000 (11/2000) 19, 19 (10/00)			
94	11/2000 (11/2000) 19, 19 (10/00)			
93	11/2000 (11/2000) 19, 19 (10/00)			
92	11/2000 (11/2000) 19, 19 (10/00)			
91	11/2000 (11/2000) 19, 19 (10/00)			
90	11/2000 (11/2000) 19, 19 (10/00)			
89	11/2000 (11/2000) 19, 19 (10/00)			
88	11/2000 (11/2000) 19, 19 (10/00)			
87	11/2000 (11/2000) 19, 19 (10/00)			
86	11/2000 (11/2000) 19, 19 (10/00)			
85	11/2000 (11/2000) 19, 19 (10/00)			
84	11/2000 (11/2000) 19, 19 (10/00)			
83	11/2000 (11/2000) 19, 19 (10/00)			
82	11/2000 (11/2000) 19, 19 (10/00)			
81	11/2000 (11/2000) 19, 19 (10/00)			
80	11/2000 (11/2000) 19, 19 (10/00)			
79	11/2000 (11/2000) 19, 19 (10/00)			
78	11/2000 (11/2000) 19, 19 (10/00)			
77	11/2000 (11/2000) 19, 19 (10/00)			
76	11/2000 (11/2000) 19, 19 (10/00)			
75	11/2000 (11/2000) 19, 19 (10/00)			
74	11/2000 (11/2000) 19, 19 (10/00)			
73	11/2000 (11/2000) 19, 19 (10/00)			
72	11/2000 (11/2000) 19, 19 (10/00)			
71	11/2000 (11/2000) 19, 19 (10/00)			
70	11/2000 (11/2000) 19, 19 (10/00)			
69	11/2000 (11/2000) 19, 19 (10/00)			
68	11/2000 (11/2000) 19, 19 (10/00)			
67	11/2000 (11/2000) 19, 19 (10/00)			
66	11/2000 (11/2000) 19, 19 (10/00)			
65	11/2000 (11/2000) 19, 19 (10/00)			
64	11/2000 (11/2000) 19, 19 (10/00)			
63	11/2000 (11/2000) 19, 19 (10/00)			
62	11/2000 (11/2000) 19, 19 (10/00)			
61	11/2000 (11/2000) 19, 19 (10/00)			
60	11/2000 (11/2000) 19, 19 (10/00)			
59	11/2000 (11/2000) 19, 19 (10/00)			
58	11/2000 (11/2000) 19, 19 (10/00)			
57	11/2000 (11/2000) 19, 19 (10/00)			
56	11/2000 (11/2000) 19, 19 (10/00)			
55	11/2000 (11/2000) 19, 19 (10/00)			
54	11/2000 (11/2000) 19, 19 (10/00)			
53	11/2000 (11/2000) 19, 19 (10/00)			
52	11/2000 (11/2000) 19, 19 (10/00)			
51	11/2000 (11/2000) 19, 19 (10/00)			
50	11/2000 (11/2000) 19, 19 (10/00)			
49	11/2000 (11/2000) 19, 19 (10/00)			
48	11/2000 (11/2000) 19, 19 (10/00)			
47	11/2000 (11/2000) 19, 19 (10/00)			
46	11/2000 (11/2000) 19, 19 (10/00)			
45	11/2000 (11/2000) 19, 19 (10/00)			
44	11/2000 (11/2000) 19, 19 (10/00)			
43	11/2000 (11/2000) 19, 19 (10/00)			
42	11/2000 (11/2000) 19, 19 (10/00)			
41	11/2000 (11/2000) 19, 19 (10/00)			
40	11/2000 (11/2000) 19, 19 (10/00)			
39	11/2000 (11/2000) 19, 19 (10/00)			
38	11/2000 (11/2000) 19, 19 (10/00)			
37	11/2000 (11/2000) 19, 19 (10/00)			
36	11/2000 (11/2000) 19, 19 (10/00)			
35	11/2000 (11/2000) 19, 19 (10/00)			
34	11/2000 (11/2000) 19, 19 (10/00)			
33	11/2000 (11/2000) 19, 19 (10/00)			
32	11/2000 (11/2000) 19, 19 (10/00)			
31	11/2000 (11/2000) 19, 19 (10/00)			
30	11/2000 (11/2000) 19, 19 (10/00)			
29	11/2000 (11/2000) 19, 19 (10/00)			
28	11/2000 (11/2000) 19, 19 (10/00)			
27	11/2000 (11/2000) 19, 19 (10/00)			
26	11/2000 (11/2000) 19, 19 (10/00)			
25	11/2000 (11/2000) 19, 19 (10/00)			
24	11/2000 (11/2000) 19, 19 (10/00)			
23	11/2000 (11/2000) 19, 19 (10/00)			
22	11/2000 (11/2000) 19, 19 (10/00)			
21	11/2000 (11/2000) 19, 19 (10/00)			
20	11/2000 (11/2000) 19, 19 (10/00)			
19	11/2000 (11/2000) 19, 19 (10/00)			
18	11/2000 (11/2000) 19, 19 (10/00)			
17	11/2000 (11/2000) 19, 19 (10/00)			
16	11/2000 (11/2000) 19, 19 (10/00)			
15	11/2000 (11/2000) 19, 19 (10/00)			
14	11/2000 (11/2000) 19, 19 (10/00)			
13	11/2000 (11/2000) 19, 19 (10/00)			
12	11/2000 (11/2000) 19, 19 (10/00)			
11	11/2000 (11/2000) 19, 19 (10/00)			
10	11/2000 (11/2000) 19, 19 (10/00)			
9	11/2000 (11/2000) 19, 19 (10/00)			
8	11/2000 (11/2000) 19, 19 (10/00)			
7	11/2000 (11/2000) 19, 19 (10/00)			
6	11/2000 (11/2000) 19, 19 (10/00)			
5	11/2000 (11/2000) 19, 19 (10/00)			
4	11/2000 (11/2000) 19, 19 (10/00)			
3	11/2000 (11/2000) 19, 19 (10/00)			
2	11/2000 (11/2000) 19, 19 (10/00)			
1	11/2000 (11/2000) 19, 19 (10/00)			

Q9UGP4
ID Q9UGP4 PRELIMINARY: PRT: 9 AA.
AC Q9UGP4;
DT 01-MAY-2000 (TEMBUREL, 13, Created)
DT 01-MAY-2000 (TEMBUREL, 13, Last sequence update)
DT 01-MAY-2000 (TEMBUREL, 13, Last annotation update)
DE D134105.1 (PROMINER OF GENE IN SPOHNCE AL023513) (FRAGMENT).
GN SE61.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SPOHNCE FROM N.A.
RA Lloyd D.;
RL Submitted (09C-1996) to the EMBL/GenBank/TrEMBL databases
DR EMBL; AL078460; CMB51751.1;
FT NON_TER
FT 1
SQ SEQUENCE 9 AA: 1125 MW: 8154A2GH05H0411 CRC64;

Query Match 1.0%; Score 4; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 TREY 82
DB 2 TREY 5

RESULT 13
Q9AXH8
ID Q9AXH8 PRELIMINARY: PRT: 9 AA.
AC Q9AXH8;
DT 01-JUN-2001 (TEMBUREL, 17, Created)
DT 01-JUN-2001 (TEMBUREL, 17, Last sequence update)
DT 01-JUN-2001 (TEMBUREL, 17, Last annotation update)
DE TIP AQUAPORIN (FRAGMENT).
GN MIPK.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicotyledons;
OC Caryophyllales; Caryophyllales; Alcegastraceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN 11
RP SEQUENCE FROM N.A.
RA Fennett J.F., Goughley F.L., Bolwell R.P.;
RT "The promoter of mipk gene which encodes for a TIP aquaporin from Mesembryanthemum crystallinum."
RL Submitted (09C-2000) to the EMBL/GenBank/TrEMBL databases.
DR EMBL; AF126737; AAC349597.1;
KW Porin.
FT NON_TER
FT 9
SQ SEQUENCE 9 AA: 969 MW: C3ED0879CDD7287D CRC64;

Query Match 1.0%; Score 4; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 247 GLAF 250
DB 3 GLAF 6

RESULT 14
Q66545
ID Q66545 PRELIMINARY: PRT: 9 AA.
AC Q66545;
DT 01-NOV-1996 (TEMBUREL, 01, Created)
DT 01-NOV-1996 (TEMBUREL, 01, Last sequence update)
DT 01-NOV-1998 (TEMBUREL, 08, Last annotation update)
DE EPILOPE C14 (FRAGMENT).
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gammathepesvirinae; Gammaherpesviruses.
OX NCBI_TaxID=10376;
RN 11
RP SEQUENCE FROM N.A.
RC STEVAIN-B95-8;
FX MERRILLN-89296424; PubMed: 2841116;
KA Wells D., Gannon P.;
RT "The expression of novel antigens from the Epstein-Barr virus large internal repeat."
RL EMBL J. 7:1191-1196(1989).
DR EMBL; X07817; CAA40676.1;
FT NON_TER
FT 1
SQ SEQUENCE 9 AA: 929 MW: 82740P10P75625 CRC64;

Query Match 1.0%; Score 4; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 PSL5 225
DB 1 PSL5 4

RESULT 15
P82938
ID P82938 PRELIMINARY: PRT: 10 AA.
AC P82938;
DT 01-MAR-2001 (TEMBUREL, 16, Created)
DT 01-MAR-2001 (TEMBUREL, 16, Last sequence update)
DT 01-JUN-2001 (TEMBUREL, 17, Last annotation update)
DE HERR-KR EXPRESSION LEVELS C (FRAGMENT).
OS Bordet vulgare (barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticaceae; Hordeum.
OX NCBI_TaxID=4513;
RN 11
RP SEQUENCE.
RC STEVAIN-B95-8; EMBL: X07817; CAA40676.1;
RX MERRILLN-89296424; PubMed: 2841116;
SA Fennett J.F., Goughley F.L., Bolwell R.P.;
RT "Separation and characterization of basic barley seed proteins."
RL Electrophoresis 21:4693-4700(2000).
CC 1; B13DELANDEK3; ON THE 20-CE; THE DEFINITIVE F1 F THIS WORK;
CC PROTEIN IS 8.5-9.0; ITS MW IS 11.9 KDA.
FT NON_TER
FT 10
SQ SEQUENCE 10 AA: 1053 MW: 9F629C30A672AAE CRC64;

Query Match 1.0%; Score 4; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 GCDG 243
DB 2 GCDG 5

RESULT 16
Q94787
ID Q94787 PRELIMINARY: PRT: 10 AA.
AC Q94787;
DT 01-DEC-2001 (TEMBUREL, 19, Created)
DT 01-DEC-2001 (TEMBUREL, 19, Last sequence update)
DT 01-DEC-2001 (TEMBUREL, 19, Last annotation update)
DE SK11 POTATOESIDE CHANNE (FRAGMENT).
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicotyledons;
OC Asteridae; eustroide I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN 11
RP SEQUENCE FROM N.A.

[illegible]

007	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035	036	037	038	039	040	041	042	043	044	045	046	047	048	049	050	051	052	053	054	055	056	057	058	059	060	061	062	063	064	065	066	067	068	069	070	071	072	073	074	075	076	077	078	079	080	081	082	083	084	085	086	087	088	089	090	091	092	093	094	095	096	097	098	099	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible][illegible]

University Match 1-6; Scores 4-10-11; Length 200
 Post Length 10,000; Prod. No. 00001
 Matchbox 4; Volume 1; Matchbox 1; Length 10,000; Prod. No. 00001

[illegible][illegible][illegible][illegible]

RESULT 21

Q47029 PRELIMINARY: PRT: 7 AA.
 AC 047029;
 DT 01-NOV-1996 (TEMBUREL 01, last sequence update)
 DT 01-NOV-1996 (TEMBUREL 01, last sequence update)
 DT 01-NOV-1996 (TEMBUREL 01, last sequence update)
 DE AAD A1 PROTEIN (FRAGMENT).
 GN AAD A1.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Enterobacter.
 OX NCBI_TaxID=550;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE 9407949; PubMed-9257126;
 RA Kather P N, Mann P A, Mierzwa P, Hare P S, Miller C H, Shaw K J;
 RT "Analysis of the *aac(7)*-*Via* gene encoding a novel 3-N-
 acetyltransferase."
 PL Antimicrob Agents Chemother 47:2074-2079(1993)
 DP FMBL: M88012; AAA15193.1;
 FT NON_TER 1
 SO SEQUENCE 7 AA: 744 MW: 6336200331A030 CP064;

Query Match 0.7%; Score 3; DR 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 336 TKV 338
 DB 2 TKV 4

RESULT 22

050556 PRELIMINARY: PRT: 7 AA.
 AC 050556;
 DT 01-JUN-1998 (TEMBUREL 06, created)
 DT 01-JUN-1998 (TEMBUREL 06, last sequence update)
 DT 01-JUN-2001 (TEMBUREL 17, last annotation update)
 DE GUYA (FRAGMENT).
 GN GUYA.
 OS Actinobacillus actinomycetococcoides (Hemophilus
 OC Actinomycetococcoides).
 OC Bacteri; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_TaxID=714;
 RN 111
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33384;
 RA MEDLINE 96355846; PubMed-8751884;
 RA Knecht G E, Spitznagel T V, Wang P, Phillips H, Jacobs C,
 RT "cis Elements and trans factors are both important in strain-specific
 RT regulation of the leukotoxin gene in Actinobacillus
 RT actinomycetococcoides."
 RL Infect. Immun. 64:3451-3460(1996).
 DP FMBL: U51862; AAB8721.1;
 FT NON_TER 1
 SO SEQUENCE 7 AA: 832 MW: 4092420767449420 CP054;

Query Match 0.7%; Score 3; DR 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 PUY 394
 DB 4 PUY 6

RESULT 23

P70804

ID P70804 PRELIMINARY: PRT: 7 AA.
 AC P70804;
 DT 01-FEB-1997 (TEMBUREL 02, created)
 DT 01-FEB-1997 (TEMBUREL 02, last sequence update)
 DT 01-DEC-2001 (TEMBUREL 19, last annotation update)
 DE ALCT PROTEIN (FRAGMENT).
 GN ALCT.
 OS Acetobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; pseudomonadaceae;
 OC Acetobacter.
 OX NCBI_TaxID=354;
 RN 111
 RP SEQUENCE FROM N.A.
 RX STRAIN=J;
 RX MEDLINE 96427318; PubMed 8840682;
 RA Rehm B H A, Ertelova H, Valla S;
 FT "A new Acetobacter vinelandii mannuronan C-5 epimerase gene (*alc1*) is
 FT part of an *alc* gene cluster physically organized in a manner similar
 FT to that in *Pseudomonas aeruginosa*."
 RL J. Bacteriol. 178:5884-5889(1996).
 DP FMBL: X87973; CAA1240.1;
 FT NON_TER 1
 SO SEQUENCE 7 AA: 684 MW: 7185A5A41A1A10 CP064;

Query Match 0.7%; Score 3; DR 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 IVS 55
 DB 2 IVS 4

RESULT 24

054248 PRELIMINARY: PRT: 7 AA.
 AC 054248;
 DT 01-NOV-1996 (TEMBUREL 01, created)
 DT 01-NOV-1996 (TEMBUREL 01, last sequence update)
 DT 01-DEC-2001 (TEMBUREL 19, last annotation update)
 DE RPLD PROTEIN (FRAGMENT).
 GN RPLD.
 OS Streptomyces griseus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetaceae; Streptomyces; Streptomyetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN 111
 RP SEQUENCE FROM N.A.
 RX STRAIN=N2-3-11;
 RX MEDLINE 20011291; PubMed 10642330;
 RA Pochling S, Piepersberg W, Wobmer U F;
 RT "Analysis and regulation of the *roc Y* gene from Streptomyces griseus
 RT N2-3-11 and relevance of the *roc Y* gene product."
 RL Biochim. Biophys. Acta 1117:299-302(1999).
 DP FMBL: X95415; CAA65160.1;
 FT NON_TER 1
 SO SEQUENCE 7 AA: 760 MW: 720720182182A0 CP064;

Query Match 0.7%; Score 3; DR 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 VIV 54
 DB 1 VIV 4

RESULT 25

P93233 PRELIMINARY: PRT: 7 AA.
 AC P93233;
 DT 01-MAY-1997 (TEMBUREL 03, created)
 DT 01-MAY-1997 (TEMBUREL 03, last sequence update)

RP SEQUENCE:
 RX MEDLINE-95290767; PubMed 7772836;
 RA Nakayama J., Abe Y., Ono Y., Isegai A., Suzuki A.;
 RI "Isolation and structure of the Enterococcus faecalis sex pheromone,
 RT cphI, that induces conjugal transfer of the hemolysin/bacteriocin
 RI plasmids, pGB1 and pYII."
 RL Riosci Biotechnol Biochem. 59:703-705(1995).
 SO SEQUENCE 8 AA: 741 MW: 830877320722002 CRC64;

Query Match 0.78; Score 3; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5,6e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 343 VLV 345
 DB 3 VLV 5

RESULT 30

ID 09P517 PRELIMINARY; PRT; 8 AA.
 AC 09P517;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-NOV-2000 (TREMBLrel. 14, Last annotation update)
 DE 1,4-BETA-D-GLUCAN GLUCANHYDROLASE (EC 3.2.1.4) (FRAGMENT)
 OS Clostridium thermocellum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1515;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-9231950; PubMed-1567379;
 RA Romanec M.P., Fauch O., Kohayashi T., Huskisson N.S., Parker P.D.,
 RA Demain A.L.;
 RI "Purification and characterization of a new endoglucanase from
 RT Clostridium thermocellum."
 RL Biochem. J. 283:69-73(1992).
 SO SEQUENCE 8 AA: 823 MW: C2C1A81DD9D1B775 CRC64;

Query Match 0.78; Score 3; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5,6e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 268 FAE 270
 DB 4 FAE 6

RESULT 31

ID 09X3K1 PRELIMINARY; PRT; 8 AA.
 AC 09X3K1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE CYTOCHROME B (FRAGMENT)
 CN PFTB.
 OS Prochlorococcus sp.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcales;
 OC Prochlorococcus.
 OX NCBI_TaxID=1220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Urbach E., Chisholm S.W.;
 RI "Genetic diversity in Prochlorococcus populations (flow cytometrically
 RT sorted from the Sargasso Sea and Gulf Stream)."
 RL Limnol. Oceanogr. 43:1615-1630(1998).
 DR EMBL: AF070193; AAD23233.1;
 FT NON-TER 1
 SO SEQUENCE 8 AA: 794 MW: 106706656/285663 CRC64;

Query Match 0.78; Score 3; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 5,6e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 SEP 17
 DB 5 SEP 7

RESULT 32

ID 087471 PRELIMINARY; PRT; 8 AA.
 AC 087471;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE H1PA (FRAGMENT)
 CN H1PA.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivisions; Pasteurellales;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN PACAN (H1A);
 EX MEDLINE 98389609; PubMed 9721816;
 RA Mullanba-Muhammad T., Merlin G., Smith A.L., Hirschfeld A.,
 RA Gotsch M.;
 RI "Evolution of the major filous gene cluster of Haemophilus
 RT influenzae."
 RL J. Bacteriol. 180:4694-4703(1998).
 DR EMBL: AF071762; AAC34830.1;
 FT NON-TER 1
 SO SEQUENCE 8 AA: 876 MW: 1844451A7272425 CRC64;

Query Match 0.78; Score 3; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5,6e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 INN 61
 DB 5 INN 7

RESULT 33

ID 09Z1Z9 PRELIMINARY; PRT; 8 AA.
 AC 09Z1Z9;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE 2-ISOPROPYLMALATE SYNTHASE (EC 4.1.3.12) (FRAGMENT)
 CN IFOA.
 OS Buchnera aphidicola
 OC Bacteroidetes; Bacteroidia; gamma subdivisions; Bacteri-
 OC Bacteroides.
 OX NCBI_TaxID=9;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99026934; PubMed-9812361;
 RA Silva F.J., Van Ham R.C.H.J., Sabater B., Latour A.;
 RI "Structure and evolution of the leucine plasmids carried by the
 RT aphidicola (Buchnera aphidicola) from aphids of the family
 RL Aphididae."
 RL FEMS Microbiol. Lett. 168:43-49(1998).
 DR EMBL: AJ006874; CAA67290.1;
 FT NON-TER 8
 SO SEQUENCE 8 AA: 917 MW: FE295042D658446 CRC64;

Query Match 0.78; Score 3; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5,6e-05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

OX NCBI_TaxID=280;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=610, AND CV. M50;
RA MEDLINE=96173113; PubMed=7868610;
RX Van der Pluijm J., Williams M., van Halbeek H.,
RT "Adaptation of Xanthobacter autotrophicus strain to bromoacetate due to
RT activation and mobilization of the haloacetate dehalogenase gene by
RL J. Bacteriol. 177:1348-1356(1995).
DR EMBL: X84038; CAA59857.1;
FT NON-TER
SQ
SEQUENCE 8 AA: 562 MW: 57494.02; pI=5.00; Cys=4.
Query Match
Best Local Similarity 100.0%; Score 3; DB 2; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 329 YPV 331
DB 6 YPV 8
RESULT 39
093454
ID 093454 PRELIMINARY: PRT: 8 AA.
AC 093454:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MERN PROTEIN (FPAGMENT)
GN MERN
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=666; TRANSPOSE=TN5037;
RA Kalyaeva E.S., Khododil G.Y., Bass I.A., Gorbunko A.M., Yurina G.V.,
RA Nikiforov V.G.;
RT "tns037, a Tn21 like mercury resistance transposon from Thiobacillus
RT ferrooxidans.";
RT Russ. J. Genet. 37:972-975(2001)
DR EMBL: AJ251743; CAC59252.1;
FT NON-TER
SQ
SEQUENCE 8 AA: 947 MW: 101542.77; pI=4.44; Cys=4.
Query Match
Best Local Similarity 100.0%; Score 3; DB 2; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 394 YPV 396
DB 4 YPV 6
RESULT 40
090899
ID 090899 PRELIMINARY: PRT: 8 AA.
AC 090899:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE AMINOPEPTIDASE B
OS Saccharomyces cerevisiae (Baker's yeast);
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycotaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RA MEDLINE=9208419; PubMed 1710199;
RA Kassel D.R., Williams K.P., Muschman R.L., Smith T.A.;
RT "Optimization of the fragmentation in a trifluoroacetic acid bombardment ion

```

```

RT Source for the sequencing of peptides at the peptide level.";
RA Anal. Chem. 64:1978-1984(1991).
SQ
SEQUENCE 8 AA: 772 MW: 78304.00; pI=5.00; Cys=64;
Query Match
Best Local Similarity 100.0%; Score 3; DB 3; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 343 YLV 345
DB 2 YLV 4
RESULT 41
09285
ID 09285 PRELIMINARY: PRT: 8 AA.
AC 09285:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CLOTTING FACTOR VIII (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shibata M., Shima M., Morichika S., Yoshida A.;
RT "Human clotting factor VIII gene, junction regions of the deletion of
RT exon 4 through 7.";
RT Submitted (MAR-2000) to the EMBL, GenBank, DDBJ databases.
DR EMBL: AB040872; DAA04312.1;
FT NON-TER
SQ
SEQUENCE 8 AA: 866 MW: 101698.74; pI=5.00; Cys=64;
Query Match
Best Local Similarity 100.0%; Score 3; DB 4; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 167 ASL 169
DB 2 ASL 4
RESULT 42
09HC00
ID 09HC00 PRELIMINARY: PRT: 8 AA.
AC 09HC00:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHODIESTERASE 10A7 (PDE10A7) (FPAGMENT).
GN HSPDE10A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujishige K., Koyota J., Yuasa K., Omori K.;
RT "The human phosphodiesterase PDE10A gene: genomic organization and
RT evolutionary relatedness with other PDEs containing GAF domains.";
RT Hum. Mol. Genet. 10:1411-1419(2001).
DR EMBL: AB041175; BAB15368.1;
FT NON-TER
SQ
SEQUENCE 8 AA: 960 MW: 104819.56; pI=5.00; Cys=64;
Query Match
Best Local Similarity 100.0%; Score 3; DB 4; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


DT 01-DEC-2001 (TREMUREL, 19, Last sequence update)
 DT 01-DEC-2001 (TREMUREL, 19, Last sequence update)
 DT MITOCHONDRIAL RIBOSOMAL PROTEIN L39 (PEACHENI)
 GN MRPL39.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
 CC NCBI_TaxID=9606;
 RN [1]
 PP SEQUENCE FROM N.A.
 RX MEDLINE=21429115; PubMed=11543634;
 RA Kameuchi N, Suzuki T, Tsuchi T, Majumori M, Yanita M, Hara S,
 RA Matsumoto K, Tanaka T.
 PT "The human mitochondrial ribosomal protein genes: Mapping of 54 genes
 PT to the chromosomes and implications for human disorders."
 RL Genomics 77:65-70(2001)
 DR EMBL: AB051346; BAR54936.1;
 KW Ribosomal protein.
 FT NON_TER
 FT NON_TER
 SO SEQUENCE 8 AA; 875 MW; 47BB5BH9CD086580 CRO64;
 Query Match 0.7%; Score 3; DR 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 OY 14 ISG 16
 DB 1 ISG 3
 RESULT 48
 ID 015896 PRELIMINARY; PRT; 8 AA.
 AC 015896;
 DT 01-JAN-1998 (TREMUREL, 05, Created)
 DT 01-JAN-1998 (TREMUREL, 05, Last sequence update)
 DT 01-DEC-2001 (TREMUREL, 19, Last annotation update)
 GN 1203 ANTIGEN (FRAGMENT).
 GN 1203.
 OS Babesia bovis.
 CC Eukaryota; Alveolata; Apicomplexa; Filizoplasmida; Babesia;
 CC NCBI_TaxID=5865;
 RN [1]
 PP SEQUENCE FROM N.A.
 RX STRAIN=SAMPD01 ATPINATERY;
 RA Silins G.U., Blakeley P.L., Riddles P.W.;
 PT "Characterization of the transcriptional control region of the 1203
 PT antigen gene from the sporozoan Babesia bovis."
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases
 DR EMBL: U44917; AAB65362.1;
 FT NON_TER
 FT NON_TER
 SO SEQUENCE 8 AA; 984 MW; PDD979G41AD929 GPC64;
 Query Match 0.7%; Score 3; DR 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 OY 78 ATP 80
 DB 3 ATP 5
 RESULT 49
 ID P82686 PRELIMINARY; PRT; 8 AA.

AC P82686;
 DT 01-MAR-2001 (TREMUREL, 16, Created)
 DT 01-MAR-2001 (TREMUREL, 16, Last sequence update)
 DT 01-MAR-2001 (TREMUREL, 16, Last sequence update)
 GN KININ-2 (PEA-K-2).
 OS Periplaneta americana (American cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 CC Phryganea; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 CC Blattodea; Blattella; Periplaneta.
 CC NCBI_TaxID=6978;
 RN [1]
 PP SEQUENCE, AND FUNCTIONAL
 RX TISSUE=COELOPODA CAROTINACEA;
 EX PROTEIN 92010462; PubMed 9450979;
 RA Prodel P., Kellner R., Rapus J., Penzlin H., Gaede G.;
 PT "Isolation and structural elucidation of eight kinins from the
 PT retrocaval complex of the American cockroach, Periplaneta
 PT americana."
 RL Regul. Pept. 71:199-205(1997).
 CC -1- FUNCTION: METALLO-PROTEINASES AND KININACTIVITY
 CC -1- SIMILARITY: HELPNES TO THE KININ FAMILY.
 CC (MYOEPHIC ACTIVITY).
 CC Neuropeptide: Amidation.
 FT MID_RES
 FT MID_RES
 SO SEQUENCE 8 AA; 656 MW; D6555A5B5A555A 7B344;
 Query Match 0.7%; Score 3; DR 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 OY 363 SWG 365
 DB 6 SWG 8
 RESULT 50
 ID P82688 PRELIMINARY; PRT; 8 AA.
 AC P82688;
 DT 01-MAR-2001 (TREMUREL, 16, Created)
 DT 01-MAR-2001 (TREMUREL, 16, Last sequence update)
 DT 01-MAR-2001 (TREMUREL, 16, Last annotation update)
 GN KININ-4 (PEA-K-4).
 OS Periplaneta americana (American cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 CC Phryganea; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 CC Blattodea; Blattella; Periplaneta.
 CC NCBI_TaxID=6978;
 RN [1]
 PP SEQUENCE, AND FUNCTIONAL
 RX TISSUE=COELOPODA CAROTINACEA;
 RX PRODEL 9450979;
 RA Prodel P., Kellner R., Rapus J., Penzlin H., Gaede G.;
 PT "Isolation and structural elucidation of eight kinins from the
 PT retrocaval complex of the American cockroach, Periplaneta
 PT americana."
 RL Regul. Pept. 71:199-205(1997).
 CC -1- FUNCTION: METALLO-PROTEINASES AND KININACTIVITY
 CC -1- SIMILARITY: HELPNES TO THE KININ FAMILY.
 CC (MYOEPHIC ACTIVITY).
 CC Neuropeptide: Amidation.
 FT MID_RES
 FT MID_RES
 SO SEQUENCE 8 AA; 656 MW; D6555A5B5A555A 7B344;
 Query Match 0.7%; Score 3; DR 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
 OY 363 SWG 365
 DB 6 SWG 8

Search completed successfully. No results found.
10/9/2002 3:14:08 PM

GenCorp version 6.1.2
Copyright (c) 1993 - 2002 Compton Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002 12:43:41 : Search time 19 seconds

(without alignments)
515,509 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 401

Sequence: 1 MDTGNVFPQPIISCPYIA.....KSTCGILVEPWYVINGSI 401

Scoring table: OLIGO

Gapop 60.0 : Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 70601

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA:
1: US-09-598-062-4
2: US-09-598-062-4
3: US-09-598-062-4
4: US-09-598-062-4
5: US-09-598-062-4
6: US-09-598-062-4

Print No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	1.7	7	4	US-08-827-1718-13
2	5	1.2	5	4	US-08-827-1718-15
3	5	1.2	5	4	US-07-752-101A-9
4	5	1.2	5	4	US-08-093-741-72
5	5	1.2	5	4	US-08-486-721A-19
6	5	1.2	5	4	US-08-720-012-72
7	5	1.2	5	4	US-07-052-754B-41
8	5	1.2	5	4	US-08-093-741-69
9	5	1.2	5	4	US-08-486-721A-5
10	5	1.2	5	4	US-08-486-721A-10
11	5	1.2	5	4	US-08-486-721A-10
12	5	1.2	5	4	US-08-486-721A-10
13	5	1.2	5	4	US-08-160-604-116
14	5	1.2	5	4	US-08-160-604-117
15	5	1.2	5	4	US-08-160-604-118
16	5	1.2	5	4	US-08-338-634-37
17	5	1.2	5	4	US-08-454-207A-36
18	5	1.2	5	4	US-08-454-207A-36
19	5	1.2	5	4	US-08-440-179-19
20	5	1.2	5	4	US-08-440-179-19
21	5	1.2	5	4	US-08-454-207A-36
22	5	1.2	5	4	US-08-454-207A-36
23	5	1.2	5	4	US-08-454-207A-36
24	5	1.2	5	4	US-08-454-207A-36
25	5	1.2	5	4	US-08-454-207A-36
26	5	1.2	5	4	US-08-454-207A-36
27	5	1.2	5	4	US-08-454-207A-36
28	5	1.2	5	4	US-08-454-207A-36

28	4	1.0	4	3	US-09-166-094-7	Sequence 7, Appl
29	4	1.0	4	3	US-09-166-094-7	Sequence 7, Appl
30	4	1.0	4	3	US-09-166-094-7	Sequence 7, Appl
31	4	1.0	4	4	US-09-246-160-54	Sequence 14, Appl
32	4	1.0	4	6	552-089-9	Patent No. 552-089
33	4	1.0	4	5	US-08-456-840-34	Sequence 44, Appl
34	4	1.0	4	5	US-08-340-428A-46	Sequence 47, Appl
35	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
36	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
37	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
38	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
39	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
40	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
41	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
42	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
43	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
44	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
45	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
46	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
47	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
48	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
49	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
50	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
51	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
52	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
53	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
54	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
55	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
56	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
57	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
58	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
59	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
60	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
61	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
62	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
63	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
64	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
65	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
66	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
67	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
68	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
69	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
70	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
71	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
72	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
73	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
74	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
75	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
76	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
77	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
78	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
79	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
80	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
81	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
82	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
83	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
84	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
85	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
86	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
87	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
88	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
89	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
90	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
91	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
92	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
93	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
94	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
95	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
96	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
97	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
98	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
99	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl
100	4	1.0	4	5	US-08-456-840-34	Sequence 47, Appl

539	4	1.0	9	US-08-765-641-3	Sequence 3, April	612	4	5	PCT-US93-08366-20	Sequence 20, April
540	4	1.0	9	US-08-470-335-25	Sequence 25, April	613	4	5	PCT-US93-08366-25	Sequence 25, April
541	4	1.0	9	US-08-612-473-93	Sequence 93, April	614	4	5	PCT-US93-08366-93	Sequence 93, April
542	4	1.0	9	US-08-789-334F-34	Sequence 34F, April	615	4	5	PCT-US93-08366-34	Sequence 34F, April
543	4	1.0	9	US-08-753-425A-347	Sequence 347, April	616	4	5	PCT-US93-08366-347	Sequence 347, April
544	4	1.0	9	US-08-268-754-123	Sequence 123, April	617	4	5	PCT-US93-08366-123	Sequence 123, April
545	4	1.0	9	US-09-258-754-370	Sequence 370, April	618	4	5	PCT-US93-08366-370	Sequence 370, April
546	4	1.0	9	US-08-358-754-327	Sequence 327, April	619	4	5	PCT-US93-08366-327	Sequence 327, April
547	4	1.0	9	US-08-158-802-175	Sequence 175, April	620	4	5	PCT-US93-08366-175	Sequence 175, April
548	4	1.0	9	US-08-169-915-12	Sequence 12, April	621	4	5	PCT-US93-08366-12	Sequence 12, April
549	4	1.0	9	US-08-202-832-9	Sequence 9, April	622	4	5	PCT-US93-08366-9	Sequence 9, April
550	4	1.0	9	US-08-725-221-25	Sequence 25, April	623	4	5	PCT-US93-08366-25	Sequence 25, April
551	4	1.0	9	US-08-731-664A-25	Sequence 25, April	624	4	5	PCT-US93-08366-25	Sequence 25, April
552	4	1.0	9	US-08-660-092-18	Sequence 18, April	625	4	5	PCT-US93-08366-18	Sequence 18, April
553	4	1.0	9	US-08-378-313-1	Sequence 1, April	626	4	5	PCT-US93-08366-1	Sequence 1, April
554	4	1.0	9	US-08-112-796-497	Sequence 497, April	627	4	5	PCT-US93-08366-497	Sequence 497, April
555	4	1.0	9	US-08-704-344-6	Sequence 6, April	628	4	5	PCT-US93-08366-6	Sequence 6, April
556	4	1.0	9	US-08-393-273E-18	Sequence 18, April	629	4	5	PCT-US93-08366-18	Sequence 18, April
557	4	1.0	9	US-08-478-373-35	Sequence 35, April	630	4	5	PCT-US93-08366-35	Sequence 35, April
558	4	1.0	9	US-09-842-197-123	Sequence 123, April	631	4	5	PCT-US93-08366-123	Sequence 123, April
559	4	1.0	9	US-09-842-197-229	Sequence 229, April	632	4	5	PCT-US93-08366-229	Sequence 229, April
560	4	1.0	9	US-09-842-197-377	Sequence 377, April	633	4	5	PCT-US93-08366-377	Sequence 377, April
561	4	1.0	9	US-08-927-597-93	Sequence 93, April	634	4	5	PCT-US93-08366-93	Sequence 93, April
562	4	1.0	9	US-08-773-156-30	Sequence 30, April	635	4	5	PCT-US93-08366-30	Sequence 30, April
563	4	1.0	9	US-08-126-171-50	Sequence 50, April	636	4	5	PCT-US93-08366-50	Sequence 50, April
564	4	1.0	9	US-09-510-738A-59	Sequence 59, April	637	4	5	PCT-US93-08366-59	Sequence 59, April
565	4	1.0	9	US-09-510-738A-111	Sequence 111, April	638	4	5	PCT-US93-08366-111	Sequence 111, April
566	4	1.0	9	US-09-510-738A-123	Sequence 123, April	639	4	5	PCT-US93-08366-123	Sequence 123, April
567	4	1.0	9	US-08-510-738A-128	Sequence 128, April	640	4	5	PCT-US93-08366-128	Sequence 128, April
568	4	1.0	9	US-08-510-738A-147	Sequence 147, April	641	4	5	PCT-US93-08366-147	Sequence 147, April
569	4	1.0	9	US-08-510-738A-160	Sequence 160, April	642	4	5	PCT-US93-08366-160	Sequence 160, April
570	4	1.0	9	US-09-133-094F-31	Sequence 31, April	643	4	5	PCT-US93-08366-31	Sequence 31, April
571	4	1.0	9	US-09-214-094F-27	Sequence 27, April	644	4	5	PCT-US93-08366-27	Sequence 27, April
572	4	1.0	9	US-09-135-839F-37	Sequence 37, April	645	4	5	PCT-US93-08366-37	Sequence 37, April
573	4	1.0	9	US-09-105-839F-42	Sequence 42, April	646	4	5	PCT-US93-08366-42	Sequence 42, April
574	4	1.0	9	US-08-105-839F-58	Sequence 58, April	647	4	5	PCT-US93-08366-58	Sequence 58, April
575	4	1.0	9	US-08-518-046-54	Sequence 54, April	648	4	5	PCT-US93-08366-54	Sequence 54, April
576	4	1.0	9	US-08-518-046-56	Sequence 56, April	649	4	5	PCT-US93-08366-56	Sequence 56, April
577	4	1.0	9	US-08-518-046-50	Sequence 50, April	650	4	5	PCT-US93-08366-50	Sequence 50, April
578	4	1.0	9	US-08-518-046-77	Sequence 77, April	651	4	5	PCT-US93-08366-77	Sequence 77, April
579	4	1.0	9	US-08-518-046-93	Sequence 93, April	652	4	5	PCT-US93-08366-93	Sequence 93, April
580	4	1.0	9	US-08-518-046-150	Sequence 150, April	653	4	5	PCT-US93-08366-150	Sequence 150, April
581	4	1.0	9	US-08-518-046-151	Sequence 151, April	654	4	5	PCT-US93-08366-151	Sequence 151, April
582	4	1.0	9	US-08-406-781-23	Sequence 23, April	655	4	5	PCT-US93-08366-23	Sequence 23, April
583	4	1.0	9	US-08-406-781-24	Sequence 24, April	656	4	5	PCT-US93-08366-24	Sequence 24, April
584	4	1.0	9	US-09-406-781-25	Sequence 25, April	657	4	5	PCT-US93-08366-25	Sequence 25, April
585	4	1.0	9	US-09-026-034-45	Sequence 45, April	658	4	5	PCT-US93-08366-45	Sequence 45, April
586	4	1.0	9	US-08-997-802-10	Sequence 10, April	659	4	5	PCT-US93-08366-10	Sequence 10, April
587	4	1.0	9	US-08-997-802-11	Sequence 11, April	660	4	5	PCT-US93-08366-11	Sequence 11, April
588	4	1.0	9	US-08-115-737-108	Sequence 108, April	661	4	5	PCT-US93-08366-108	Sequence 108, April
589	4	1.0	9	US-08-187-859-503	Sequence 503, April	662	4	5	PCT-US93-08366-503	Sequence 503, April
590	4	1.0	9	US-08-187-859-1251	Sequence 1251, April	663	4	5	PCT-US93-08366-1251	Sequence 1251, April
591	4	1.0	9	US-08-187-859-1676	Sequence 1676, April	664	4	5	PCT-US93-08366-1676	Sequence 1676, April
592	4	1.0	9	US-08-187-859-1706	Sequence 1706, April	665	4	5	PCT-US93-08366-1706	Sequence 1706, April
593	4	1.0	9	US-08-187-859-3074	Sequence 3074, April	666	4	5	PCT-US93-08366-3074	Sequence 3074, April
594	4	1.0	9	US-08-187-859-3075	Sequence 3075, April	667	4	5	PCT-US93-08366-3075	Sequence 3075, April
595	4	1.0	9	US-08-187-859-3078	Sequence 3078, April	668	4	5	PCT-US93-08366-3078	Sequence 3078, April
596	4	1.0	9	US-08-187-859-3119	Sequence 3119, April	669	4	5	PCT-US93-08366-3119	Sequence 3119, April
597	4	1.0	9	US-08-187-859-3121	Sequence 3121, April	670	4	5	PCT-US93-08366-3121	Sequence 3121, April
598	4	1.0	9	US-08-187-859-3123	Sequence 3123, April	671	4	5	PCT-US93-08366-3123	Sequence 3123, April
599	4	1.0	9	US-08-187-859-3164	Sequence 3164, April	672	4	5	PCT-US93-08366-3164	Sequence 3164, April
600	4	1.0	9	US-08-187-859-3165	Sequence 3165, April	673	4	5	PCT-US93-08366-3165	Sequence 3165, April
601	4	1.0	9	US-08-187-859-3166	Sequence 3166, April	674	4	5	PCT-US93-08366-3166	Sequence 3166, April
602	4	1.0	9	US-08-187-859-3167	Sequence 3167, April	675	4	5	PCT-US93-08366-3167	Sequence 3167, April
603	4	1.0	9	US-08-187-859-3168	Sequence 3168, April	676	4	5	PCT-US93-08366-3168	Sequence 3168, April
604	4	1.0	9	US-08-187-859-3211	Sequence 3211, April	677	4	5	PCT-US93-08366-3211	Sequence 3211, April
605	4	1.0	9	US-08-187-859-3213	Sequence 3213, April	678	4	5	PCT-US93-08366-3213	Sequence 3213, April
606	4	1.0	9	US-08-187-859-3254	Sequence 3254, April	679	4	5	PCT-US93-08366-3254	Sequence 3254, April
607	4	1.0	9	US-08-187-859-3255	Sequence 3255, April	680	4	5	PCT-US93-08366-3255	Sequence 3255, April
608	4	1.0	9	US-08-187-859-3256	Sequence 3256, April	681	4	5	PCT-US93-08366-3256	Sequence 3256, April
609	4	1.0	9	US-08-187-859-3257	Sequence 3257, April	682	4	5	PCT-US93-08366-3257	Sequence 3257, April
610	4	1.0	9	US-08-187-859-3258	Sequence 3258, April	683	4	5	PCT-US93-08366-3258	Sequence 3258, April
611	4	1.0	9	US-08-187-859-3259	Sequence 3259, April	684	4	5	PCT-US93-08366-3259	Sequence 3259, April

841	1	0	10	5	PCT-US96-03189-26	Sequence 26, App1	904	3	0.7	4	1	US-08-432-617-6	Sequence 5, App1
842	4	1	0	5	PCT-US96-03189-26	Sequence 29, App1	905	3	0.7	4	1	US-08-432-617-6	Sequence 5, App1
843	4	1	0	5	PCT-US96-03189-34	Sequence 34, App1	906	3	0.7	4	1	US-08-243-982-27	Sequence 24, App1
834	4	1	0	5	PCT-US96-03189-35	Sequence 35, App1	907	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
835	4	1	0	5	PCT-US96-03189-36	Sequence 36, App1	908	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
836	4	1	0	5	PCT-US96-03189-37	Sequence 37, App1	909	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
837	4	1	0	5	PCT-US96-03189-38	Sequence 38, App1	910	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
838	4	1	0	5	PCT-US96-06270-134	Sequence 134, App1	911	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
839	4	1	0	5	PCT-US96-06270-138	Sequence 138, App1	912	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
840	4	1	0	5	PCT-US96-06270-142	Sequence 142, App1	913	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
841	4	1	0	5	PCT-US96-06270-146	Sequence 146, App1	914	3	0.7	4	1	US-08-014-979-11	Sequence 7, App1
842	4	1	0	5	PCT-US96-11945-5	Sequence 5, App1	915	3	0.7	4	1	US-08-149-849-20	Sequence 20, App1
843	4	1	0	6	5169933-18	Patient No. 5169933	916	3	0.7	4	1	US-08-190-111-1	Sequence 11, App1
844	4	1	0	6	5169933-41	Patient No. 5169933	917	3	0.7	4	1	US-07-972-007-26	Sequence 26, App1
845	4	1	0	6	5221735-4	Patient No. 5221735	918	3	0.7	4	1	US-08-167-346A-1	Sequence 167, App1
846	4	1	0	6	5252228-2	Patient No. 5252228	919	3	0.7	4	1	US-08-213-195A-19	Sequence 19, App1
847	4	1	0	6	526545-22	Patient No. 526545	920	3	0.7	4	1	US-08-327-682-16	Sequence 16, App1
848	3	0.7	3	1	US-07-815-670A-8	Sequence 8, App1	921	3	0.7	4	1	US-08-140-188-3	Sequence 3, App1
849	3	0.7	3	1	US-08-124-544A-21	Sequence 21, App1	922	3	0.7	4	1	US-08-213-195A-19	Sequence 19, App1
850	3	0.7	3	1	US-08-384-618-1	Sequence 1, App1	923	3	0.7	4	1	US-08-210-766A-17	Sequence 17, App1
851	3	0.7	3	1	US-08-167-336A-3	Sequence 3, App1	924	3	0.7	4	1	US-07-991-1995-1	Sequence 1, App1
852	3	0.7	2	1	US-08-116-762-3	Sequence 3, App1	925	3	0.7	4	1	US-07-973-135-1	Sequence 1, App1
853	3	0.7	3	1	US-08-589-011-3	Sequence 3, App1	926	3	0.7	4	1	US-08-544-342-4	Sequence 4, App1
854	3	0.7	3	1	US-07-923-724-44	Sequence 44, App1	927	3	0.7	4	1	US-07-872-6746-7	Sequence 7, App1
855	3	0.7	3	1	US-08-843-035-24	Sequence 24, App1	928	3	0.7	4	1	US-07-872-6746-7	Sequence 7, App1
856	3	0.7	3	1	US-08-689-426A-44	Sequence 44, App1	929	3	0.7	4	1	US-08-378-106-13	Sequence 13, App1
857	3	0.7	3	2	US-08-755-406A-24	Sequence 24, App1	930	3	0.7	4	1	US-08-208-108-14	Sequence 14, App1
858	3	0.7	3	2	US-08-948-762-3	Sequence 3, App1	931	3	0.7	4	1	US-08-431-539-2	Sequence 2, App1
859	3	0.7	3	2	US-08-451-822A-14	Sequence 14, App1	932	3	0.7	4	1	US-08-021-608B-24	Sequence 23, App1
860	3	0.7	3	2	US-08-544-432-14	Sequence 14, App1	933	3	0.7	4	1	US-08-123-762-4	Sequence 4, App1
861	3	0.7	3	2	US-08-595-047A-19	Sequence 19, App1	934	3	0.7	4	1	US-08-357-430-4	Sequence 4, App1
862	3	0.7	3	2	US-08-601-455A-9	Sequence 9, App1	935	3	0.7	4	1	US-08-211-070A-7	Sequence 7, App1
863	3	0.7	3	2	US-09-060-455-9	Sequence 9, App1	936	3	0.7	4	1	US-08-411-412-71	Sequence 71, App1
864	3	0.7	3	2	US-08-464-025A-18	Sequence 18, App1	937	3	0.7	4	1	US-08-076-092-27	Sequence 27, App1
865	3	0.7	3	2	US-08-464-025A-23	Sequence 23, App1	938	3	0.7	4	1	US-08-076-092-27	Sequence 27, App1
866	3	0.7	3	4	US-09-083-630-9	Sequence 9, App1	939	3	0.7	4	1	US-08-076-092-41	Sequence 41, App1
867	3	0.7	3	4	US-09-461-697-310	Sequence 310, App1	940	3	0.7	4	1	US-08-243-873-1	Sequence 1, App1
868	3	0.7	3	5	PCT-US94-11270-8	Sequence 8, App1	941	3	0.7	4	1	US-08-376-941-6	Sequence 6, App1
869	3	0.7	3	5	PCT-US94-09143-21	Sequence 21, App1	942	3	0.7	4	1	US-08-201-081-1	Sequence 1, App1
870	3	0.7	4	1	US-09-595-081A-4	Sequence 4, App1	943	3	0.7	4	1	US-08-201-081-1	Sequence 1, App1
871	3	0.7	4	1	US-07-670-706-8	Sequence 8, App1	944	3	0.7	4	1	US-08-249-297-1	Sequence 297, App1
872	3	0.7	4	1	US-07-822-924-6	Sequence 6, App1	945	3	0.7	4	1	US-08-451-682-20	Sequence 20, App1
873	3	0.7	4	1	US-07-811-048-4	Sequence 4, App1	946	3	0.7	4	1	US-08-240-712-6	Sequence 6, App1
874	3	0.7	4	1	US-07-772-087-7	Sequence 7, App1	947	3	0.7	4	1	US-07-890-4256-77	Sequence 77, App1
875	3	0.7	4	1	US-07-862-021B-6	Sequence 6, App1	948	3	0.7	4	1	US-08-331-383-7	Sequence 7, App1
876	3	0.7	4	1	US-07-862-021B-8	Sequence 8, App1	949	3	0.7	4	1	US-08-331-383-9	Sequence 9, App1
877	3	0.7	4	1	US-07-725-431-57	Sequence 57, App1	950	3	0.7	4	1	US-08-331-383-21	Sequence 21, App1
878	3	0.7	4	1	US-07-801-812A-5	Sequence 5, App1	951	3	0.7	4	1	US-08-331-383-22	Sequence 22, App1
879	3	0.7	4	1	US-07-932-200-4	Sequence 4, App1	952	3	0.7	4	1	US-07-969-605-61	Sequence 61, App1
880	3	0.7	4	1	US-07-973-852-4	Sequence 4, App1	953	3	0.7	4	1	US-07-969-605-62	Sequence 62, App1
881	3	0.7	4	1	US-07-820-154A-20	Sequence 20, App1	954	3	0.7	4	1	US-08-462-697-5	Sequence 5, App1
882	3	0.7	4	1	US-08-093-781-9	Sequence 9, App1	955	3	0.7	4	1	US-08-213-195A-19	Sequence 19, App1
883	3	0.7	4	1	US-07-791-2140-21	Sequence 21, App1	956	3	0.7	4	1	US-08-471-052A-17	Sequence 17, App1
884	3	0.7	4	1	US-07-950-773-4	Sequence 4, App1	957	3	0.7	4	1	US-08-335-198-17	Sequence 17, App1
885	3	0.7	4	1	US-08-203-716-24	Sequence 24, App1	958	3	0.7	4	1	US-08-335-198-25	Sequence 25, App1
886	3	0.7	4	1	US-07-841-997A-5	Sequence 5, App1	959	3	0.7	4	1	US-08-335-198-31	Sequence 31, App1
887	3	0.7	4	1	US-07-930-649-8	Sequence 8, App1	960	3	0.7	4	1	US-08-321-625-81	Sequence 81, App1
888	3	0.7	4	1	US-07-805-727-36	Sequence 36, App1	961	3	0.7	4	1	US-08-280-444-39	Sequence 39, App1
889	3	0.7	4	1	US-08-106-504-7	Sequence 7, App1	962	3	0.7	4	1	US-08-544-144A-8	Sequence 8, App1
890	3	0.7	4	1	US-07-594-983-2	Sequence 2, App1	963	3	0.7	4	1	US-08-458-998A-15	Sequence 15, App1
891	3	0.7	4	1	US-07-694-983-4	Sequence 4, App1	964	3	0.7	4	1	US-08-458-998A-4	Sequence 4, App1
892	3	0.7	4	1	US-07-828-450-33	Sequence 33, App1	965	3	0.7	4	1	US-08-458-998A-7	Sequence 7, App1
893	3	0.7	4	1	US-08-384-618-3	Sequence 3, App1	966	3	0.7	4	1	US-08-458-998A-9	Sequence 9, App1
894	3	0.7	4	1	US-08-384-618-6	Sequence 6, App1	967	3	0.7	4	1	US-08-458-998A-12	Sequence 12, App1
895	3	0.7	4	1	US-08-384-618-8	Sequence 8, App1	968	3	0.7	4	1	US-08-458-998A-16	Sequence 16, App1
896	3	0.7	4	1	US-07-972-387-21	Sequence 21, App1	969	3	0.7	4	1	US-08-458-998A-19	Sequence 19, App1
897	3	0.7	4	1	US-07-969-407A-5	Sequence 5, App1	970	3	0.7	4	1	US-08-458-998A-19	Sequence 19, App1
898	3	0.7	4	1	US-08-127-904-2	Sequence 2, App1	971	3	0.7	4	1	US-08-358-140-108	Sequence 108, App1
899	3	0.7	4	1	US-08-127-904-3	Sequence 3, App1	972	3	0.7	4	1	US-08-358-140-108	Sequence 108, App1
900	3	0.7	4	1	US-08-127-904-3	Sequence 3, App1	973	3	0.7	4	1	US-08-358-140-108	Sequence 108, App1
901	3	0.7	4	1	US-08-210-742A-1	Sequence 1, App1	974	3	0.7	4	1	US-08-432-617-6	Sequence 6, App1
902	3	0.7	4	1	US-08-174-365A-61	Sequence 61, App1	975	3	0.7	4	1	US-08-142-430A-8	Sequence 8, App1
903	3	0.7	4	1	US-08-174-365A-105	Sequence 105, App1	976	3	0.7	4	1	US-08-142-430A-8	Sequence 8, App1

COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/09/093,741
FILING DATE: 20-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 443 23 754.1
FILING DATE: 15-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELEPHONE: (202)628-8803
TELEFAX: (202)628-8244
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-093-741-69

Query Match 1.2%, Score 5, 18 1: length 7:
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 PKAY 290
DB 2 PKAY 6

RESULT 9
US-08-486-721A-5
Sequence 5, Application 08/09486721A
Patent No. 5749025
GENERAL INFORMATION:
APPLICANT: Fukazawa, Chikafusa
TITLE OF INVENTION: Method of Producing
TITLE OF INVENTION: Aspartaglyl Endoprotease
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frishaut, Holitz, Goodman, Langer & Chick
STREET: 767 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10017-2023
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 4 inch, 1.4 mb
COMPUTER: IBM PC
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/09/486,721A
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/091,991
FILING DATE: 12-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Barth, Richard S.
REGISTRATION NUMBER: 28,180
REFERENCE/DOCKET NUMBER: 940897/103
TELEPHONE: (212) 319-4400
TELEFAX: (212) 319-5101
TELEX: 236268

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHEetical:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANISM:
TEMPERATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
FOCUS/STAIN INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
US-08-486-721A-5

Query Match 1.2%, Score 5, 18 1: length 7:
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NIVE 9
DB 2 NIVE 6

RESULT 10
US-08-486-721A-10
Sequence 10, Application 08/09486721A
Patent No. 5749025
GENERAL INFORMATION:
APPLICANT: Fukazawa, Chikafusa
TITLE OF INVENTION: Method of Producing
TITLE OF INVENTION: Aspartaglyl Endoprotease
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frishaut, Holitz, Goodman, Langer & Chick
STREET: 767 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10017-2023
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 4 inch, 1.4 mb
COMPUTER: IBM PC
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/09/486,721A
FILING DATE: 7-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:


```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 319-4980
TELEFAX: (312) 319-5101
TELEX: 236268
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
US-08-486-721A-6

Query Match      1.28; Score 5; DB 1; Length 8;
Best Local Similarity 100.0%; P-adj No. 1 7e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 5 NNVEE 9
DB 3 NNVEE 7

RESULT 13
US-08-160-604-116
Sequence 116 Application US/08160604
Patent No. 623522
GENERAL INFORMATION:
APPLICANT: Harley, John
APPLICANT: James, Judith A.
APPLICANT: Scofield, R. H.
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc-dos/ms-dos
SOFTWARE: Patent In Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATOR ROBBIN, US/08160,604
FILING DATE: 30 NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,319
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DECKET NUMBER: 08N11431E(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-160-604-116

Query Match      1.28; Score 5; DB 4; Length 8;
Best Local Similarity 100.0%; P-adj No. 1 7e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 80 REYR 84
DB 3 REYR 7

RESULT 14
US-08-160-604-117
Sequence 117 Application US/08160604
Patent No. 623522
GENERAL INFORMATION:
APPLICANT: Harley, John
APPLICANT: James, Judith A.
APPLICANT: Scofield, R. H.
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc-dos/ms-dos
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,604
FILING DATE: 30-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,819
FILING DATE: 13 APR 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,205
FILING DATE: 31-JAN-1991

```


RESULT 17
US-08-454-207A-36
Sequence 36, Application US/08454207A
Patent No. 5710123
GENERAL INFORMATION
APPLICANT: Heaven, George A.
APPLICANT: Kruszynski, Marian
TITLE OF INVENTION: PEPTIDE INHIBITORS OF SELECTIN BINDING
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESS: Washington, D.C. 20001-1111
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,207A
FILING DATE: 09-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P91/0594,12119
FILING DATE: 13-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/997,771
FILING DATE: 18-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Dianne B. Elderklin
REGISTRATION NUMBER: 28,598
REFERENCE/AGENT NUMBER: 000R-0183
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-584-4100
TELEFAX: 215-568-9439
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-454-207A-36

Query Match 1 24; Score 5; DP 1; Length 9;
Best local similarity 100.0%; Ident. No. 1; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 375 KLAH 379
IIIIII
DB 4 KLAH 8

RESULT 18
US-08-283-716-19
Sequence 19, Application US/08283716
Patent No. 5416013
GENERAL INFORMATION:
APPLICANT: BLACK, ROY A
APPLICANT: SLEATH, PAUL R
APPLICANT: KRONHEIM, SHIRLEY R
TITLE OF INVENTION: INTERLEUKIN 1B PROTEASE AND
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: PESSIER, JILL SMITH, SHOPP, SUTTER &
STREET: 180 N. STETSON
CITY: CHICAGO
STATE: IL
COUNTRY: USA
ZIP: 60601

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,716
FILING DATE:
CLASSIFICATION: 415
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05/077,504,411
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KATE, MARTIN L.
REGISTRATION NUMBER: 25011
REFERENCE/AGENT NUMBER: 000RINX2108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165400
TELEFAX: 3126165460
TELEX: 9102211206
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-203-716-19

Query Match 1 24; Score 5; DP 1; Length 10;
Best local similarity 100.0%; Ident. No. 2; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 329 VPADP 333
IIIIII

DB 3 VPADP 7

RESULT 19
US-08-440-179-19
Sequence 19, Application US/08440179
Patent No. 5756465
GENERAL INFORMATION:
APPLICANT: BLACK, ROY A
APPLICANT: SLEATH, PAUL R
APPLICANT: KRONHEIM, SHIRLEY R
TITLE OF INVENTION: INTERLEUKIN 1B PROTEASE AND
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: PESSIER, JILL SMITH, SHOPP, SUTTER &
STREET: 180 N. STETSON
CITY: CHICAGO
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,179
FILING DATE: 12-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05-07750,644
FILING DATE: 30-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: KATZ, MARTIN L.
REGISTRATION NUMBER: 25011
REFERENCE/AGENT NUMBER: 000RINX2108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165400

US-09-336-447A-89

Query Match 1.28; Score 5; DB 4, Length 10,
Best Local Similarity 100.0%; Prod. No. 2, 10-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NPIPE 9
DB 4 NPIPE 8

RESULT 23

US-08-361-862-4

Sequence 4, Application US/08361862
Patent No. 5725964
GENERAL INFORMATION:
APPLICANT: Yamamoto, Naoki
APPLICANT: Nakashima, Hideki
APPLICANT: Morisuchi, Wataru
APPLICANT: Tanaka, Shigeki
APPLICANT: Dosako, Shunichi
APPLICANT: Shimoto, Hiroshi
TITLE OF INVENTION: A composition for enhancing infection
TITLE OF INVENTION: and growth of human immunodeficiency virus
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Patent Administrator, Testa, Hurwitz &
ADDRESSER: Thibault
STREET: 53 State Street, Exchange Place
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcom Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08361862
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/097,848
FILING DATE: 27-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Paula A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FIN-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7000
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-361-862-4

Query Match 1.0%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Prod. No. 1, 70-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 GRKG 384
DB 1 GRKG 4

RESULT 24

US-08-623-833B-67
Sequence 67, Application US/08623833B

Patent No. 5866684

GENERAL INFORMATION:
APPLICANT: SHIMURA, Ryoohito
APPLICANT: KASAI, Konichi
APPLICANT: MATSUMOTO, Shiroaki
APPLICANT: TAKAMOTO, Hisayoshi
TITLE OF INVENTION: ISOLEUCINE POINT MARKERS FOR
TITLE OF INVENTION: ELECTRIC TRAVELING WITH THE SEQUENCE LIPETIN
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Pillsbury, Madison & Sutro, L.L.C.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-2918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/0556333B
FILING DATE: 29-MAR-1996
CLASSIFICATION: 540
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9798471995
FILING DATE: 41-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 271166/1995
FILING DATE: 19-OCT-1995
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-623-833B-67

Query Match 1.0%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Prod. No. 1, 70-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 UNDER 158
DB 1 UNDER 4

RESULT 25

US-08-392-438A-7

Sequence 7, Application US/08392438A
Patent No. 5869620
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James F.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filippa, David
TITLE OF INVENTION: Multivalent Antigen Binding Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSER: Stretto, Kossler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

RESULT 29
 US-09-166-093-7
 : Sequence 7, Application us/09166094
 : Patent No. 6027725
 : GENERAL INFORMATION:
 : APPLICANT: WHITLOW, Marc
 : APPLICANT: Wood, James F.
 : APPLICANT: Hardman, Karl
 : APPLICANT: Bird, Robert
 : APPLICANT: Filipula, David
 : APPLICANT: Kellence, Michelle
 : TITLE OF INVENTION: Multivalent Antigen Binding Proteins
 : NUMBER OF SEQUENCES: 23
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sterne, Kossler, Goldstein & Fox P.L.L.C.
 : STREET: 1100 New York Avenue, NW
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20005
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/166,093
 : FILING DATE: Herewith
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/392,338
 : FILING DATE: 22-FEB-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/989,846
 : FILING DATE: 20-NOV-1992
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/796,976
 : FILING DATE: 25-NOV-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Goldstein, Jorge A.
 : REGISTRATION NUMBER: 29,021
 : REFERENCE/AGENT NUMBER: 0977,000000
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 371-2540
 : TELEFAX: (202) 371-2540
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4 amino acids
 : TYPE: amino acid
 : TOPOLOGY: both
 : US-09-166-093-7
 : Query Match 1.00; Score 4; DB 3; Length 4;
 : Best Local Similarity 100.00; Pred. No. 1.7e+05;
 : Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Q7 52 VTVS 55
 ||||
 DB 1 VTVS 4

RESULT 29
 US-09-172-019-7
 : Sequence 7, Application us/09172019
 : Patent No. 6103889
 : GENERAL INFORMATION:
 : APPLICANT: WHITLOW, Marc
 : APPLICANT: Hardman, Karl
 : APPLICANT: Bird, Robert
 : APPLICANT: Filipula, David
 : TITLE OF INVENTION: Nucleic Acid Molecules Encoding Signal Chain

TITLE OF INVENTION: Antigen-Binding Proteins (AS Amended)
 : NUMBER OF SEQUENCES: 23
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sterne, Kossler, Goldstein & Fox P.L.L.C.
 : STREET: 1100 New York Avenue, NW
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20005
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/172,019
 : FILING DATE: Herewith
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/492,348
 : FILING DATE: 22-FEB-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/989,846
 : FILING DATE: 20-NOV-1992
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/796,976
 : FILING DATE: 25-NOV-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Goldstein, Jorge A.
 : REGISTRATION NUMBER: 29,021
 : REFERENCE/AGENT NUMBER: 0977,000000
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 371-2540
 : TELEFAX: (202) 371-2540
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4 amino acids
 : TYPE: amino acid
 : TOPOLOGY: both
 : US-09-172-019-7
 : Query Match 1.00; Score 4; DB 3; Length 4;
 : Best Local Similarity 100.00; Pred. No. 1.7e+05;
 : Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Q7 52 VTVS 55
 ||||
 DB 1 VTVS 4

RESULT 30
 US-09-166-094-7
 : Sequence 7, Application us/09166094
 : Patent No. 6121424
 : GENERAL INFORMATION:
 : APPLICANT: WHITLOW, Marc
 : APPLICANT: Wood, James F.
 : APPLICANT: Hardman, Karl
 : APPLICANT: Bird, Robert
 : APPLICANT: Filipula, David
 : APPLICANT: Kellence, Michelle
 : TITLE OF INVENTION: Multivalent Antigen Binding Proteins
 : NUMBER OF SEQUENCES: 23
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sterne, Kossler, Goldstein & Fox P.L.L.C.
 : STREET: 1100 New York Avenue, NW
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20005
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible

RESULT 34

US-08-340-4288-46

Sequence 46, Application US/08/464288

Patent No. 5648465

GENERAL INFORMATION:

APPLICANT: MARGOLIS, Richard U.

APPLICANT: RAUCH, Ivo

APPLICANT: MARGOLIS, Renee K.

TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A

TITLE OF INVENTION: NEURKAN AS A CHONDROITIN SULFATE PROTEOGLYCAN

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Browdy and Neimark

STREET: 419 Seventh Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,4288

FILING DATE: 14 No. 5648465ember 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 87/922,711

FILING DATE: 03 August 1992

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.

REGISTRATION NUMBER: 95,619

REFERENCE/DOCKET NUMBER: Margolis 1A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-340-4288-46

Query Match 1.0%; Score 4; DB 1; Length 5;

Best Local Similarity 100.0%; Prod. No. 1,760,05;

Matches 4; Conservative 0; Mismatches 0; Indels 0;

DB 2 1SCP 5

US-08-211-747-6

Sequence 6, Application US/08/211747

Patent No. 5734549

GENERAL INFORMATION:

APPLICANT: Yamada, Shingo

APPLICANT: Inoue, Keiichi

APPLICANT: Kitajima, Megumi

APPLICANT: Sawarabayashi, Kunosuke

TITLE OF INVENTION: Peptides Including Amino Acid Sequences

TITLE OF INVENTION: Selected From Hypotensin (a) and Angiotensin (a).

TITLE OF INVENTION: Antibodies Recognizing These Amino Acid Sequences, and

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMUNDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10039-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 87/922,711

FILING DATE: 14-Apr-1994

CLASSIFICATION: 540

ATTORNEY/AGENT INFORMATION:

NAME: Mastro, S. Leslie

REGISTRATION NUMBER: 7005-077

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 790-8647/741

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-211-747-6

Query Match 1.0%; Score 4; DB 1; Length 5;

Best Local Similarity 100.0%; Prod. No. 1,760,05;

Matches 4; Conservative 0; Mismatches 0; Indels 0;

DB 1 AVAA 4

US-08-486-721A-20

Sequence 20, Application US/08/486721A

Patent No. 5739025

GENERAL INFORMATION:

APPLICANT: Fukazawa, Chikafusa

TITLE OF INVENTION: Method of Producing

TITLE OF INVENTION: Asparticyl Endoprotease

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Frisbair, Holtz, Goodman, Lander & Chick

STREET: 767 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10017-2024

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 4 inch, 1.4 mb

COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 87/038,746

FILING DATE: 7-JUNE-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 82/091,991

FILING DATE: 12-JULY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Barth, Richard S.

REGISTRATION NUMBER: 28,180

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 679-1000
 TELEFAX: (212) 679-1001
 COUNCIL ON Foreign Relations
 1775 Avenue of the Americas
 16th Floor
 New York, NY 10019-1186
 Tel: (212) 679-1000

Country: USA
 Postal Code: 10019
 City: New York
 State: NY

Web: www.cfr.org

DB: 1 NAME: 1

RECORD: 1
 RECORD NO: 1

RECORD NO: 1
 RECORD NO: 1

GENERAL INFORMATION

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

DB: 1 NAME: 1

RECORD: 1

RECORD NO: 1

RECORD NO: 1

GENERAL INFORMATION

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

APPELLANT: JAMES P. GILLESPIE

STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,407A
FILING DATE: 27-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,441
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-258-5200
TELEFAX: 301-977-0847
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-266-407A-34

Query Match 1.08; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KSR 164
DB 1 KSR 4

RESULT 40
US-08-665-220-25
Sequence 25, Affiliation 09/09/94/5200
Patent No. 5786173
GENERAL INFORMATION:
APPLICANT: Almeida, Enad S.
APPLICANT: Fernandes-Almeida, Teresa
APPLICANT: Litwack, Gerald
APPLICANT: Armstrong, Robert
APPLICANT: Tomaselli, Kevin
TITLE OF INVENTION: Meth and Meth5, Apoptotic Proteases,
TITLE OF INVENTION: Nucleic Acids Encoding and Methods of Use
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESS: Campbell & Flores LLP
STREET: 4470 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/565,220
FILING DATE: 14-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/618,408
FILING DATE: 19-MAR-1996
ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Kathryn A.
REGISTRATION NUMBER: 31,815
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 545-9001
TELEFAX: (619) 545-8949
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1-5
OTHER SEQ INFORMATION: None
US-08-665-220-25

Query Match 1.08; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 315 LOAD 418
DB 1 LOAD 4

RESULT 41
US-08-437-462-19
Sequence 19, Application 05/06/97/362
Patent No. 5811104 5776469
GENERAL INFORMATION:
APPLICANT: Dale, Beverly
APPLICANT: Yamashita, Miles
APPLICANT: Acree, William M.
APPLICANT: Chavez Jr., Lloyd O.
TITLE OF INVENTION: FILTRATION METHODS FOR TREATING VIBROS
TITLE OF INVENTION: DIAGNOSTIC TOOLS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: KIRKSTON & FRIESTER
STREET: 1000 17th St NW, Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/473,362
FILING DATE: 09 MAY 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,904
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
FILING DATE: 09/04/95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/473,362
FILING DATE: 09 MAY 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,904
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
FILING DATE: 09/04/95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/473,362
FILING DATE: 09 MAY 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

Query Match 1.08; Score 4; DB 1; Length 5;

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/850,392
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/469,362
 FILING DATE: 06-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: DUVALL, JEAN M.
 REGISTRATION NUMBER: 32,731
 REFERENCE/DOCKET NUMBER: 002010-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 877-0900
 TELEFAX: (415) 877-8370
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 5
 OTHER INFORMATION: /note "This position is Tyr or
 OTHER INFORMATION: Trp."
 US-08-850-392-10

Query Match 1.00; Score 4; PR 2; Length 5;
 Best Local Similarity 100.00; Pred. No. 1.7e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 197 CASC 209
 1111
 DB 1 CASC 4

RESULT 45
 US-08-852-544-34
 Sequence: 34; Affiliation: 197-209-2544
 Patent No. 5874544
 GENERAL INFORMATION:
 APPLICANT: Taddei-Peters, W. C.
 APPLICANT: Butler, Sandra M.
 TITLE OF INVENTION: Immunoreactive peptides of Apo(a)
 NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: AKZO NO. 5874544et
 STREET: 1330 Piccard drive
 CITY: Rockville
 STATE: Maryland
 COUNTRY: US
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/892,544
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/256,497
 FILING DATE: 27-JUN-1994
 APPLICATION NUMBER: US 08/172,461
 FILING DATE: 21-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Gormley, Mary E.
 REGISTRATION NUMBER: 34,409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-258-5200
 TELEFAX: 301-977-0847
 INFORMATION FOR SEQ ID NO: 34:
 1.00; Score 4; PR 2; Length 5;
 Best Local Similarity 100.00; Pred. No. 1.7e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 STRANDEDNESS: single
 US-08-892-544-34

Query Match 1.00; Score 4; PR 2; Length 5;
 Best Local Similarity 100.00; Pred. No. 1.7e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 341 KSC 164
 1111
 DB 1 KSSR 4

RESULT 46
 US-08-647-7598-81
 Sequence 81; Application US/08/6477598
 Patent No. 5876941
 GENERAL INFORMATION:
 APPLICANT: David William Holden
 TITLE OF INVENTION: Identification of genes
 NUMBER OF SEQUENCES: 501
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patton H. Fabst
 STREET: 2800 rue Atlantic Road
 STREET: 1201 West Peachtree Street
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30329-4450
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/647,7598
 FILING DATE: 03 MAY 1996
 CLASSIFICATION: A45
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/02875
 FILING DATE: 11 DEC 1995
 CLASSIFICATION: A45
 ATTORNEY/AGENT INFORMATION:
 NAME: Patist, Patricia L.
 REGISTRATION NUMBER: 41,284
 REFERENCE/DOCKET NUMBER: P008111
 TELEPHONE: (404) 876-8794
 TELEFAX: (404) 876-8796
 INFORMATION FOR SEQ ID NO: 81:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYDROPHILIC: NO
 US-08-647-7598-81

Query Match 1.00; Score 4; PR 2; Length 5;
 Best Local Similarity 100.00; Pred. No. 1.7e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 248 LAFU 251
 1111
 DB 2 LAFU 5

RESULT 47
 US-08-928-958-20


```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24 MAR 1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB91/03121
FILING DATE: 10-JUL-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00605
FILING DATE: 24-MAR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/150,002
FILING DATE: 31-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/307,619
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: CLOUGH, DAVID W
REGISTRATION NUMBER: 46,107
REFERENCE/CHECK NUMBER: 28111/32372
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
INFORMATION FOR SEQ ID NO: 318:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
US-08-350-260A-318

```

```

Query Match          1.0%  Score 4; 108 2; Length 5;
Best Local Similarity 100.0%  Pred. No. 1,76+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 77 PATR 80
DB 1 PAIR 4

```

```

RESULT 50
US-08-755-728-6
Sequence 6, Application US/08755728
Patent No. 5962312
GENERAL INFORMATION:
APPLICANT: Ploman, Gregory
APPLICANT: Mossie, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF ADR-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSO for windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09-755,728
FILING DATE: NO. 5962312 October 25, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/008,809
FILING DATE: December 18, 1995
APPLICATION NUMBER: 60/023,943

```

```

FILING DATE: August 14, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 423427
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 455-0440
TELEX: 673510
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-09-755-728-6

```

```

Query Match          1.0%  Score 4; 108 2; Length 5;
Best Local Similarity 100.0%  Pred. No. 1,76+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 ISCP 17
DB 2 ISCP 5

```

Search completed: October 9, 2002, 12:47:24
Job time : 24 secs

221	221	1.7	201	2.2	AA071904	Human gene 1, cloned
222	222	1.7	201	2.2	AA071904	Human gene 2, cloned
223	223	1.7	201	2.2	AA071904	Human gene 3, cloned
224	224	1.7	201	2.2	AA071904	Human gene 4, cloned
225	225	1.7	201	2.2	AA071904	Human gene 5, cloned
226	226	1.7	201	2.2	AA071904	Human gene 6, cloned
227	227	1.7	201	2.2	AA071904	Human gene 7, cloned
228	228	1.7	201	2.2	AA071904	Human gene 8, cloned
229	229	1.7	201	2.2	AA071904	Human gene 9, cloned
230	230	1.7	201	2.2	AA071904	Human gene 10, cloned
231	231	1.7	201	2.2	AA071904	Human gene 11, cloned
232	232	1.7	201	2.2	AA071904	Human gene 12, cloned
233	233	1.7	201	2.2	AA071904	Human gene 13, cloned
234	234	1.7	201	2.2	AA071904	Human gene 14, cloned
235	235	1.7	201	2.2	AA071904	Human gene 15, cloned
236	236	1.7	201	2.2	AA071904	Human gene 16, cloned
237	237	1.7	201	2.2	AA071904	Human gene 17, cloned
238	238	1.7	201	2.2	AA071904	Human gene 18, cloned
239	239	1.7	201	2.2	AA071904	Human gene 19, cloned
240	240	1.7	201	2.2	AA071904	Human gene 20, cloned
241	241	1.7	201	2.2	AA071904	Human gene 21, cloned
242	242	1.7	201	2.2	AA071904	Human gene 22, cloned
243	243	1.7	201	2.2	AA071904	Human gene 23, cloned
244	244	1.7	201	2.2	AA071904	Human gene 24, cloned
245	245	1.7	201	2.2	AA071904	Human gene 25, cloned
246	246	1.7	201	2.2	AA071904	Human gene 26, cloned
247	247	1.7	201	2.2	AA071904	Human gene 27, cloned
248	248	1.7	201	2.2	AA071904	Human gene 28, cloned
249	249	1.7	201	2.2	AA071904	Human gene 29, cloned
250	250	1.7	201	2.2	AA071904	Human gene 30, cloned
251	251	1.7	201	2.2	AA071904	Human gene 31, cloned
252	252	1.7	201	2.2	AA071904	Human gene 32, cloned
253	253	1.7	201	2.2	AA071904	Human gene 33, cloned
254	254	1.7	201	2.2	AA071904	Human gene 34, cloned
255	255	1.7	201	2.2	AA071904	Human gene 35, cloned
256	256	1.7	201	2.2	AA071904	Human gene 36, cloned
257	257	1.7	201	2.2	AA071904	Human gene 37, cloned
258	258	1.7	201	2.2	AA071904	Human gene 38, cloned
259	259	1.7	201	2.2	AA071904	Human gene 39, cloned
260	260	1.7	201	2.2	AA071904	Human gene 40, cloned
261	261	1.7	201	2.2	AA071904	Human gene 41, cloned
262	262	1.7	201	2.2	AA071904	Human gene 42, cloned
263	263	1.7	201	2.2	AA071904	Human gene 43, cloned
264	264	1.7	201	2.2	AA071904	Human gene 44, cloned
265	265	1.7	201	2.2	AA071904	Human gene 45, cloned
266	266	1.7	201	2.2	AA071904	Human gene 46, cloned
267	267	1.7	201	2.2	AA071904	Human gene 47, cloned
268	268	1.7	201	2.2	AA071904	Human gene 48, cloned
269	269	1.7	201	2.2	AA071904	Human gene 49, cloned
270	270	1.7	201	2.2	AA071904	Human gene 50, cloned
271	271	1.7	201	2.2	AA071904	Human gene 51, cloned
272	272	1.7	201	2.2	AA071904	Human gene 52, cloned
273	273	1.7	201	2.2	AA071904	Human gene 53, cloned
274	274	1.7	201	2.2	AA071904	Human gene 54, cloned
275	275	1.7	201	2.2	AA071904	Human gene 55, cloned
276	276	1.7	201	2.2	AA071904	Human gene 56, cloned
277	277	1.7	201	2.2	AA071904	Human gene 57, cloned
278	278	1.7	201	2.2	AA071904	Human gene 58, cloned
279	279	1.7	201	2.2	AA071904	Human gene 59, cloned
280	280	1.7	201	2.2	AA071904	Human gene 60, cloned
281	281	1.7	201	2.2	AA071904	Human gene 61, cloned
282	282	1.7	201	2.2	AA071904	Human gene 62, cloned
283	283	1.7	201	2.2	AA071904	Human gene 63, cloned
284	284	1.7	201	2.2	AA071904	Human gene 64, cloned
285	285	1.7	201	2.2	AA071904	Human gene 65, cloned
286	286	1.7	201	2.2	AA071904	Human gene 66, cloned
287	287	1.7	201	2.2	AA071904	Human gene 67, cloned
288	288	1.7	201	2.2	AA071904	Human gene 68, cloned
289	289	1.7	201	2.2	AA071904	Human gene 69, cloned
290	290	1.7	201	2.2	AA071904	Human gene 70, cloned
291	291	1.7	201	2.2	AA071904	Human gene 71, cloned
292	292	1.7	201	2.2	AA071904	Human gene 72, cloned
293	293	1.7	201	2.2	AA071904	Human gene 73, cloned
294	294	1.7	201	2.2	AA071904	Human gene 74, cloned
295	295	1.7	201	2.2	AA071904	Human gene 75, cloned
296	296	1.7	201	2.2	AA071904	Human gene 76, cloned
297	297	1.7	201	2.2	AA071904	Human gene 77, cloned
298	298	1.7	201	2.2	AA071904	Human gene 78, cloned
299	299	1.7	201	2.2	AA071904	Human gene 79, cloned
300	300	1.7	201	2.2	AA071904	Human gene 80, cloned
301	301	1.7	201	2.2	AA071904	Human gene 81, cloned
302	302	1.7	201	2.2	AA071904	Human gene 82, cloned
303	303	1.7	201	2.2	AA071904	Human gene 83, cloned

523	6	1.5	135	22	AA013042	Human polypeptide
524	6	1.5	136	21	AA007917	Arabidopsis thaliana
525	6	1.5	136	21	AA010124	Arabidopsis thaliana
526	6	1.5	136	21	AA012084	Arabidopsis thaliana
527	6	1.5	136	21	AA026219	Arabidopsis thaliana
528	6	1.5	136	21	AA033664	Arabidopsis thaliana
529	6	1.5	136	21	AA047920	Arabidopsis thaliana
530	6	1.5	137	22	AA021981	Human choroideremia
531	6	1.5	138	22	AA065092	Gene #23, associated with
532	6	1.5	139	21	AA047096	Arabidopsis thaliana
533	6	1.5	139	22	AA020572	Arabidopsis thaliana
534	6	1.5	140	21	AA008198	Protein encoded by
535	6	1.5	140	21	AA026270	Arabidopsis thaliana
536	6	1.5	140	21	AA046054	Arabidopsis thaliana
537	6	1.5	140	21	AA055032	Arabidopsis thaliana
538	6	1.5	140	22	AA012266	Human choroideremia
539	6	1.5	140	22	AA080354	Human choroideremia
540	6	1.5	140	22	AA087759	Human choroideremia
541	6	1.5	141	21	AA016197	Arabidopsis thaliana
542	6	1.5	141	22	AA017121	Arabidopsis thaliana
543	6	1.5	142	20	AA033542	B. juncea CDF homo
544	6	1.5	142	21	AA016196	Arabidopsis thaliana
545	6	1.5	142	22	AA033845	Human choroideremia
546	6	1.5	143	20	AA033541	B. juncea CDF homo
547	6	1.5	143	21	AA033151	Pinus radiata trans
548	6	1.5	143	21	AA041424	Cellulose synthase
549	6	1.5	144	21	AA061206	Arabidopsis thaliana
550	6	1.5	144	22	AA010965	Secreted human choroideremia
551	6	1.5	145	21	AA034730	Human secreted protein
552	6	1.5	145	21	AA064272	Arabidopsis thaliana
553	6	1.5	145	22	AA026869	Arabidopsis thaliana
554	6	1.5	145	22	AA063080	Arabidopsis thaliana
555	6	1.5	145	22	AA035001	Ectoparasitic flatworm
556	6	1.5	146	21	AA034799	Plant cell wall
557	6	1.5	146	21	AA009315	Arabidopsis thaliana
558	6	1.5	146	22	AA013452	Human choroideremia
559	6	1.5	147	21	AA021288	Protein of rat choroideremia
560	6	1.5	147	21	AA021289	Arabidopsis thaliana
561	6	1.5	147	21	AA016196	Arabidopsis thaliana
562	6	1.5	147	22	AA040494	Human choroideremia
563	6	1.5	148	22	AA057294	Cysteine protease
564	6	1.5	150	16	AA064904	Proteinase 37
565	6	1.5	150	22	AA058125	Proteinase 37
566	6	1.5	150	22	AA006405	Human polypeptide
567	6	1.5	151	21	AA055023	Arabidopsis thaliana
568	6	1.5	151	20	AA033553	B. oleracea CDF homo
569	6	1.5	153	22	AA040817	Human polypeptide
570	6	1.5	155	21	AA033141	Pinus radiata trans
571	6	1.5	155	21	AA009430	Arabidopsis thaliana
572	6	1.5	155	21	AA083102	574 proteinase 37
573	6	1.5	155	22	AA030018	Human choroideremia
574	6	1.5	157	18	AA030797	The tyrosinase-related
575	6	1.5	157	21	AA064271	Arabidopsis thaliana
576	6	1.5	157	21	AA005647	Arabidopsis thaliana
577	6	1.5	157	21	AA012953	Arabidopsis thaliana
578	6	1.5	157	22	AA023146	Human choroideremia
579	6	1.5	157	22	AA006406	Human choroideremia
580	6	1.5	157	22	AA020652	Human choroideremia
581	6	1.5	158	21	AA012083	Arabidopsis thaliana
582	6	1.5	158	21	AA047922	Arabidopsis thaliana
583	6	1.5	159	22	AA067764	Proteinase 37
584	6	1.5	160	20	AA019629	SHG ID NO 347 from
585	6	1.5	160	22	AA065412	Proteinase 37
586	6	1.5	161	10	AA019111	Proteinase 37
587	6	1.5	161	21	AA013580	Sequence of the human
588	6	1.5	161	22	AA070967	Human choroideremia
589	6	1.5	162	11	AA070967	Human choroideremia
590	6	1.5	162	18	AA031586	Human choroideremia
591	6	1.5	162	20	AA042792	Human choroideremia
592	6	1.5	163	6	AA050045	Sequence of the human
593	6	1.5	163	7	AA060247	Human choroideremia
594	6	1.5	163	21	AA061205	Arabidopsis thaliana
595	6	1.5	163	22	AA060247	Arabidopsis thaliana

[illegible]

815	6	1.5	288	19	AA098446	H. pylori CIPPO 542
816	6	1.5	289	20	AA098452	γ-globulin prote
817	6	1.5	290	21	AA098457	Sequence of human
818	6	1.5	291	10	AA098460	serum albumin
819	6	1.5	292	11	AA098464	Human IgG1 heavy
820	6	1.5	293	12	AA098467	Human IgG1 heavy
821	6	1.5	294	13	AA098470	Human IgG1 heavy
822	6	1.5	295	14	AA098473	Human IgG1 heavy
823	6	1.5	296	15	AA098476	Human IgG1 heavy
824	6	1.5	297	16	AA098479	Human IgG1 heavy
825	6	1.5	298	17	AA098482	Human IgG1 heavy
826	6	1.5	299	18	AA098485	Human IgG1 heavy
827	6	1.5	300	19	AA098488	Human IgG1 heavy
828	6	1.5	301	20	AA098491	Human IgG1 heavy
829	6	1.5	302	21	AA098494	Human IgG1 heavy
830	6	1.5	303	22	AA098497	Human IgG1 heavy
831	6	1.5	304	23	AA098500	Human IgG1 heavy
832	6	1.5	305	24	AA098503	Human IgG1 heavy
833	6	1.5	306	25	AA098506	Human IgG1 heavy
834	6	1.5	307	26	AA098509	Human IgG1 heavy
835	6	1.5	308	27	AA098512	Human IgG1 heavy
836	6	1.5	309	28	AA098515	Human IgG1 heavy
837	6	1.5	310	29	AA098518	Human IgG1 heavy
838	6	1.5	311	30	AA098521	Human IgG1 heavy
839	6	1.5	312	31	AA098524	Human IgG1 heavy
840	6	1.5	313	32	AA098527	Human IgG1 heavy
841	6	1.5	314	33	AA098530	Human IgG1 heavy
842	6	1.5	315	34	AA098533	Human IgG1 heavy
843	6	1.5	316	35	AA098536	Human IgG1 heavy
844	6	1.5	317	36	AA098539	Human IgG1 heavy
845	6	1.5	318	37	AA098542	Human IgG1 heavy
846	6	1.5	319	38	AA098545	Human IgG1 heavy
847	6	1.5	320	39	AA098548	Human IgG1 heavy
848	6	1.5	321	40	AA098551	Human IgG1 heavy
849	6	1.5	322	41	AA098554	Human IgG1 heavy
850	6	1.5	323	42	AA098557	Human IgG1 heavy
851	6	1.5	324	43	AA098560	Human IgG1 heavy
852	6	1.5	325	44	AA098563	Human IgG1 heavy
853	6	1.5	326	45	AA098566	Human IgG1 heavy
854	6	1.5	327	46	AA098569	Human IgG1 heavy
855	6	1.5	328	47	AA098572	Human IgG1 heavy
856	6	1.5	329	48	AA098575	Human IgG1 heavy
857	6	1.5	330	49	AA098578	Human IgG1 heavy
858	6	1.5	331	50	AA098581	Human IgG1 heavy
859	6	1.5	332	51	AA098584	Human IgG1 heavy
860	6	1.5	333	52	AA098587	Human IgG1 heavy
861	6	1.5	334	53	AA098590	Human IgG1 heavy
862	6	1.5	335	54	AA098593	Human IgG1 heavy
863	6	1.5	336	55	AA098596	Human IgG1 heavy
864	6	1.5	337	56	AA098599	Human IgG1 heavy
865	6	1.5	338	57	AA098602	Human IgG1 heavy
866	6	1.5	339	58	AA098605	Human IgG1 heavy
867	6	1.5	340	59	AA098608	Human IgG1 heavy
868	6	1.5	341	60	AA098611	Human IgG1 heavy
869	6	1.5	342	61	AA098614	Human IgG1 heavy
870	6	1.5	343	62	AA098617	Human IgG1 heavy
871	6	1.5	344	63	AA098620	Human IgG1 heavy
872	6	1.5	345	64	AA098623	Human IgG1 heavy
873	6	1.5	346	65	AA098626	Human IgG1 heavy
874	6	1.5	347	66	AA098629	Human IgG1 heavy
875	6	1.5	348	67	AA098632	Human IgG1 heavy
876	6	1.5	349	68	AA098635	Human IgG1 heavy
877	6	1.5	350	69	AA098638	Human IgG1 heavy
878	6	1.5	351	70	AA098641	Human IgG1 heavy
879	6	1.5	352	71	AA098644	Human IgG1 heavy
880	6	1.5	353	72	AA098647	Human IgG1 heavy
881	6	1.5	354	73	AA098650	Human IgG1 heavy
882	6	1.5	355	74	AA098653	Human IgG1 heavy
883	6	1.5	356	75	AA098656	Human IgG1 heavy
884	6	1.5	357	76	AA098659	Human IgG1 heavy
885	6	1.5	358	77	AA098662	Human IgG1 heavy
886	6	1.5	359	78	AA098665	Human IgG1 heavy
887	6	1.5	360	79	AA098668	Human IgG1 heavy

PN US6254869-B1.
 XX 03-JUL-2001.
 PD
 XX 27-MAR-1997; 97US-0827171.
 PF
 XX 27-MAR-1996; 96US-0014233.
 PR
 XX (RESC) UNIV CALIFORNIA.
 PA
 XX Petersen C, Huang J;
 PI WPI: 2001-388923/41.
 DR N-PSDB; AAF84095.
 DX
 XX
 PT New vaccines and cryptopain protein for diagnosis and treatment of
 PT Cryptosporidium species infection -
 XX
 PS Claim 5; Fig 3; 32pp; English.
 CC The invention relates to vaccines and cryptopain protein for diagnosis
 CC and treatment of Cryptosporidium species infection. Cryptopain protein
 CC can be expressed by standard recombinant methodology. The vaccines
 CC comprising the cryptopain protein or its fragments are useful for active
 CC immunization of animals and humans against Cryptosporidium infection, or
 CC for production of passive immune products in admixture with an adjuvant.
 CC Cryptopain can be used for prophylactic, therapeutic, diagnostic and
 CC detection purposes. The present sequence represents the C. parvum
 CC cryptopain mature enzyme.
 CC
 SQ Sequence 226 AA;
 Query Match 56.4%; Score 226; DB 22; Length 226;
 Best local similarity 100.0%; Prod. No. 14; 217.
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 176 PNSTINWEACGVNIRNKKNTSCWAFSAVAALFAAFQIMKIPSI SGGQFVCSQ 245
 DB 1 PNSTINWEACGVNIRNKKNTSCWAFSAVAALFAAFQIMKIPSI SGGQFVCSQ 60
 CY 246 NNSFGTGGTIGALFVIAIKKLYLTNDVGYFAEKKICMLSPENYELPVKAYKVER 295
 DB 61 NGNECCDGGTIGALFVIAIKKLYLTNDVGYFAEKKICMLSPENYELPVKAYKVER 120
 QY 296 RNINLKTALAKYSPISVAIQADQTPPOPKSKVFNAPGCKRNHNVLEVMDEDTK 355
 DB 121 RNINLKTALAKYSPISVAIQADQTPPOPKSKVFNAPGCKRNHNVLEVMDEDTK 180
 QY 356 FYWIVRNSWGEAWGKCYIKLALHSGKKGCGGII VHPVYVIMQSI 401
 DB 181 FYWIVRNSWGEAWGKCYIKLALHSGKKGCGGII VHPVYVIMQSI 226
 RESULT 3
 AAB85126
 ID AAB85126 standard; Protein; 175 AA.
 XX
 AC AAB85126;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE C. parvum cryptopain pro and pro region;
 XX
 KM Cryptosporidium: infectious cryptopain and paracain; vasodilat
 KM immunization.
 XX
 OS Cryptosporidium parvum.
 XX
 PN US6254869-B1.
 PD 03-JUL-2001.
 XX
 PF 27-MAR-1997; 97US-0827171.
 PR

XX 27-MAR-1996; 96US-0014233.
 PR
 XX (RESC) UNIV CALIFORNIA.
 PA
 XX Petersen C, Huang J;
 PI WPI: 2001-388923/41.
 DR N-PSDB; AAF84094.
 DX
 XX
 PT New vaccines and cryptopain protein for diagnosis and treatment of
 PT Cryptosporidium species infection -
 XX
 PS Claim 4; Fig 3; 32pp; English.
 CC The invention relates to vaccines and cryptopain protein for diagnosis
 CC and treatment of Cryptosporidium species infection. Cryptopain protein
 CC can be expressed by standard recombinant methodology. The vaccines
 CC comprising the cryptopain protein or its fragments are useful for active
 CC immunization of animals and humans against Cryptosporidium infection, or
 CC for production of passive immune products in admixture with an adjuvant.
 CC Cryptopain can be used for prophylactic, therapeutic, diagnostic and
 CC detection purposes. The present sequence represents the pro and pro
 CC regions of C. parvum cryptopain protein.
 CC
 SQ Sequence 175 AA;
 Query Match 44.5%; Score 175; DB 22; Length 175;
 Best local similarity 100.0%; Prod. No. 14; 166.
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLCIRGRRERFVIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIR 60
 DB 1 MLCIRGRRERFVIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIR 60
 QY 61 MLCIRGRRERFVIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIR 120
 DB 61 MLCIRGRRERFVIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIRGIR 120
 QY 121 NGQTSVYVPHITGTSVFTFAATGTYIRPSYSPVYVYVYVYVYVYVYVYVYV 175
 DB 121 NGQTSVYVPHITGTSVFTFAATGTYIRPSYSPVYVYVYVYVYVYVYVYVYV 175
 RESULT 4
 AAB73748
 ID AAB73748 standard; Protein; 51 AA.
 XX
 AC AAB73748;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE cysteine protease 10 homologue;
 XX
 KM cysteine protease 10 homologue; recombinant product for;
 KM malignant tumour; cancer blood disease; HIV infection;
 KM human immunodeficiency virus; immune disorder; inflammatory condition;
 KM cystostatic; anti HIV; anti-inflammatory; immunomodulator.
 XX
 OS unidentified.
 XX
 PN W629016442 A1.
 PD 29-JUN-2001.
 XX
 PF 18 DEC-2000; 2000W-0000608.
 XX
 PP 22-JUN-1999; 99CN-0125492.
 XX
 PA (BLOW-) BLOWING GENE DEV INT CHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX

For each of these three methods, we have calculated the mean and standard deviation of the number of iterations required for convergence. The mean number of iterations required for convergence is 1.5 for the three methods. The standard deviation is 0.5 for the three methods.

$$\begin{aligned} & \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} \\ & \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{1}{\sigma_i^2} \right\}^{-1} \end{aligned}$$

Abstract

1. *Introduction*
 2. *Methodology*
 3. *Results*
 4. *Discussion*
 5. *Conclusion*
 6. *References*
 7. *Appendix*
 8. *Notes*
 9. *Tables*
 10. *Figures*
 11. *Supplementary Materials*
 12. *Author Biographies*
 13. *Conflict of Interest Statement*
 14. *Copyright*
 15. *References*
 16. *Appendix*
 17. *Notes*
 18. *Tables*
 19. *Figures*
 20. *Supplementary Materials*
 21. *Author Biographies*
 22. *Conflict of Interest Statement*
 23. *Copyright*
 24. *References*
 25. *Appendix*
 26. *Notes*
 27. *Tables*
 28. *Figures*
 29. *Supplementary Materials*
 30. *Author Biographies*
 31. *Conflict of Interest Statement*
 32. *Copyright*
 33. *References*
 34. *Appendix*
 35. *Notes*
 36. *Tables*
 37. *Figures*
 38. *Supplementary Materials*
 39. *Author Biographies*
 40. *Conflict of Interest Statement*
 41. *Copyright*
 42. *References*
 43. *Appendix*
 44. *Notes*
 45. *Tables*
 46. *Figures*
 47. *Supplementary Materials*
 48. *Author Biographies*
 49. *Conflict of Interest Statement*
 50. *Copyright*
 51. *References*
 52. *Appendix*
 53. *Notes*
 54. *Tables*
 55. *Figures*
 56. *Supplementary Materials*
 57. *Author Biographies*
 58. *Conflict of Interest Statement*
 59. *Copyright*
 60. *References*
 61. *Appendix*
 62. *Notes*
 63. *Tables*
 64. *Figures*
 65. *Supplementary Materials*
 66. *Author Biographies*
 67. *Conflict of Interest Statement*
 68. *Copyright*
 69. *References*
 70. *Appendix*
 71. *Notes*
 72. *Tables*
 73. *Figures*
 74. *Supplementary Materials*
 75. *Author Biographies*
 76. *Conflict of Interest Statement*
 77. *Copyright*
 78. *References*
 79. *Appendix*
 80. *Notes*
 81. *Tables*
 82. *Figures*
 83. *Supplementary Materials*
 84. *Author Biographies*
 85. *Conflict of Interest Statement*
 86. *Copyright*
 87. *References*
 88. *Appendix*
 89. *Notes*
 90. *Tables*
 91. *Figures*
 92. *Supplementary Materials*
 93. *Author Biographies*
 94. *Conflict of Interest Statement*
 95. *Copyright*
 96. *References*
 97. *Appendix*
 98. *Notes*
 99. *Tables*
 100. *Figures*
 101. *Supplementary Materials*
 102. *Author Biographies*
 103. *Conflict of Interest Statement*
 104. *Copyright*
 105. *References*
 106. *Appendix*
 107. *Notes*
 108. *Tables*
 109. *Figures*
 110. *Supplementary Materials*
 111. *Author Biographies*
 112. *Conflict of Interest Statement*
 113. *Copyright*
 114. *References*
 115. *Appendix*
 116. *Notes*
 117. *Tables*
 118. *Figures*
 119. *Supplementary Materials*
 120. *Author Biographies*
 121. *Conflict of Interest Statement*
 122. *Copyright*
 123. *References*
 124. *Appendix*
 125. *Notes*
 126. *Tables*
 127. *Figures*
 128. *Supplementary Materials*
 129. *Author Biographies*
 130. *Conflict of Interest Statement*
 131. *Copyright*
 132. *References*
 133. *Appendix*
 134. *Notes*
 135. *Tables*
 136. *Figures*
 137. *Supplementary Materials*
 138. *Author Biographies*
 139. *Conflict of Interest Statement*
 140. *Copyright*
 141. *References*
 142. *Appendix*
 143. *Notes*
 144. *Tables*
 145. *Figures*
 146. *Supplementary Materials*
 147. *Author Biographies*
 148. *Conflict of Interest Statement*
 149. *Copyright*
 150. *References*
 151. *Appendix*
 152. *Notes*
 153. *Tables*
 154. *Figures*
 155. *Supplementary Materials*
 156. *Author Biographies*
 157. *Conflict of Interest Statement*
 158. *Copyright*
 159. *References*
 160. *Appendix*
 161. *Notes*
 162. *Tables*
 163. *Figures*
 164. *Supplementary Materials*
 165. *Author Biographies*
 166. *Conflict of Interest Statement*
 167. *Copyright*
 168. *References*
 169. *Appendix*
 170. *Notes*
 171. *Tables*
 172. *Figures*
 173. *Supplementary Materials*
 174. *Author Biographies*
 175. *Conflict of Interest Statement*
 176. *Copyright*
 177. *References*
 178. *Appendix*
 179. *Notes*
 180. *Tables*
 181. *Figures*
 182. *Supplementary Materials*
 183. *Author Biographies*
 184. *Conflict of Interest Statement*
 185. *Copyright*
 186. *References*
 187. *Appendix*
 188. *Notes*
 189. *Tables*
 190. *Figures*
 191. *Supplementary Materials*
 192. *Author Biographies*
 193. *Conflict of Interest Statement*
 194. *Copyright*
 195. *References*
 196. *Appendix*
 197. *Notes*
 198. *Tables*
 199. *Figures*
 200. *Supplementary Materials*
 201. *Author Biographies*
 202. *Conflict of Interest Statement*
 203. *Copyright*
 204. *References*
 205. *Appendix*
 206. *Notes*
 207. *Tables*
 208. *Figures*
 209. *Supplementary Materials*
 210. *Author Biographies*
 211. *Conflict of Interest Statement*
 212. *Copyright*
 213. *References*
 214. *Appendix*
 215. *Notes*
 216. *Tables*
 217. *Figures*
 218. *Supplementary Materials*
 219. *Author Biographies*
 220. *Conflict of Interest Statement*
 221. *Copyright*
 222. *References*
 223. *Appendix*
 224. *Notes*
 225. *Tables*
 226. *Figures*
 227. *Supplementary Materials*
 228. *Author Biographies*
 229. *Conflict of Interest Statement*
 230. *Copyright*
 231. *References*
 232. *Appendix*
 233. *Notes*
 234. *Tables*
 235. *Figures*
 236. *Supplementary Materials*
 237. *Author Biographies*
 238. *Conflict of Interest Statement*
 239. *Copyright*
 240. *References*
 241. *Appendix*
 242. *Notes*
 243. *Tables*
 244. *Figures*
 245. *Supplementary Materials*
 246. *Author Biographies*
 247. *Conflict of Interest Statement*
 248

(Aldrich et al., 1986) and a number of other authors have reported that there are no significant differences between the two groups.

of "systemic feedback" for the teacher and the student, and the "systemic feedback" for the teacher and the student.

[illegible]

1. The first part of the paper is devoted to the study of the asymptotic behavior of the solutions of the system (1) as $\epsilon \rightarrow 0$. It is shown that the solutions of the system (1) converge to the solutions of the system (2) in the sense of the weak convergence in the space $L^2(\Omega; \mathbb{R}^n)$.

[illegible][illegible]

1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 26

THE UNIVERSITY OF CHICAGO

[illegible]

ANAL. Calcd for $C_{10}H_{10}O$: C, 88.10%; H, 7.41%. Found: C, 88.1%; H, 7.4%.

[illegible][illegible]

Figure 1. Schematic representation of the experimental design. The figure is divided into two main sections: 'Pre-Test' and 'Main Experiment'. The 'Pre-Test' section includes 'Pre-Test 1' (a 2x2 grid of conditions) and 'Pre-Test 2' (a 2x2 grid of conditions). The 'Main Experiment' section includes 'Main Experiment 1' (a 2x2 grid of conditions) and 'Main Experiment 2' (a 2x2 grid of conditions). The conditions are defined by the presence or absence of a 'Target' and a 'Distractor'.





































































[illegible]
$$P(A \cap B) = P(A) \cdot P(B) \quad \text{if } A \text{ and } B \text{ are independent.}$$
[illegible][illegible][illegible][illegible]

1. *Introduction*

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:01 11 November 2014

Matches 12: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAVA 208
 |||||
 DB 149 CGSCWAFSAVA 160

RESULT 7

AAW15417
 ID AAW15417 standard; Protein: 344 AA.

AC AAW15417;

DT 05-JUN-1997 (first entry)

DE EA9 protein.

KW Plant seed cell, napin gene; acyl carrier protein gene; EA9 gene;

KW Brassica, fatty acid, rat thioesterase; acyl ACP;

OS Brassica campestris.

PN US5608152-A.

PD 04-MAR-1997.

PF 31-JUL-1986; 86US-089152q

PR 09-JUL-1990; 90US-0550804

PR 31-JUL-1986; 86US-089152q

PR 28-JUL-1987; 87US-0078538.

PR 25-JAN-1988; 88US-0147781.

PR 08-APR-1988; 88US-0147781.

PR 30-MAY-1995; 95US-0453924.

PA (CALJ) CALGENE INC.

PI Knauf VC, Kridl JC;

DR WPI: 1997-14560/15

DR N-PSDB; AAT65477.

XX Transgenic Brassica plant and seeds - contain DNA construct for seed

PT specific expression of heterologous proteins

XX Example 4; Column 35-38; 48pp; English.

XX The sequences given in AAW15417 represent proteins encoded by DNA

CC fragments which were used in the construction of the DNA construct

CC of the invention. The DNA construct comprises, in the 5' to 3'

CC direction of transcription;

CC (a) a transcriptional initiation region from a gene that encodes a

CC product preferentially expressed in a plant seed cell as compared to

CC other plant cells;

CC (b) a DNA sequence of interest other than the native coding sequence

CC of the gene; and

CC (c) a transcriptional termination region.

CC Where the gene is a napin gene, an acyl carrier protein gene or an EA9

CC gene, this construct may be used in the production of a Brassica plant

CC which may be modified such that heterologous proteins are produced in

CC the seeds, such as fatty acids, rat thioesterase, acyl ACP or

CC phospholipid acyl desaturase

XX Sequence 344 AA;

XX Query Match 3.08; Score 12; DB 149; Length 344;

CC Best Local Similarity 100.0%; Pred. No. 0.0019;

CC Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 197 CGSCWAFSAVA 208

XX |||||

XX DB 149 CGSCWAFSAVA 160

RESULT 8

AAW30528
 ID AAW30528 standard; Protein: 344 AA.

AC AAW30528;

DT 26-OCT-1998 (first entry)

DE Brassica campestris seed-specific protein EA9.

KW Transgenic plant; tissue-specific gene expression;

KW seed specific expression; EA9.

OS Brassica campestris.

PN US5753475-A.

PD 19-MAY-1998.

PF 19-AUG-1993; 93US-0105852.

PR 10-AUG-1993; 93US-0105852.

PR 17-JAN-1985; 85US-0692605.

PR 31-JUL-1986; 86US-089152q.

PR 26-MAY-1987; 87US-0054369.

PR 28-JUL-1987; 87US-0078538.

PR 25-JAN-1988; 88US-0147781.

PR 15-MAY-1988; 88US-0168190.

PR 29-APR-1988; 88US-0168161.

PR 02-NOV-1988; 88US-0267855.

PR 21-MAY-1990; 90US-0526123.

PR 14-SEP-1995; 95US-0582241.

PR 08-AUG-1991; 91US-0743844.

PA (CALJ) CALGENE INC.

PI Houck CM;

DR WPI: 1998-21403/27.

DR N-PSDB; AAV36970.

XX Transformation of plants with regulatory sequence containing

XX constructs for tissue specific expression of genes

XX Example 5; Fig 4A-E; 67pp; English.

XX This polynucleotide is encoded by the coding region of a DNA of Fig

CC (see AAV36970) isolated from a Brassica campestris seed cDNA library.

CC EA9 was characterized as a seed-specific gene. The promoter region

CC of the EA9 gene can be incorporated into an expression cassette and

CC used in a novel method for producing a plant with a regulatable

CC phenotype. In this method, regulatory regions from plant genes

CC expressed during a particular developmental stage or in a specific

CC tissue are identified by cDNA screening. The regulatory regions

CC are manipulated for use with foreign sequences for introduction

CC into plant cells to provide transgenic plants having a predictable

CC phenotype that can be selected. The invention, as exemplified with

CC lightly seed and fruit specific promoters. Also claimed are methods

CC for altering the phenotype of fruit tissue as distinct to other

CC fruit tissues, modifying the genotype of a plant to impart a desired

CC characteristic to fruit, modifying transcription in fruit tissue,

CC and expressing a heterologous DNA sequence of interest specifically

CC in fruit tissue. The method provides transformation without call

CC formation in plants which have historically not been Agrobacterium

CC hosts.

XX Sequence 344 AA;

XX Query Match 3.08; Score 12; DB 149; Length 344;

XX Best Local Similarity 100.0%; Pred. No. 0.0019;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 197 CGSCWAFSAVA 208

XX |||||

XX DB 149 CGSCWAFSAVA 160

Query Match 3.0% Score 12; FR 19; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.0019;
 Matches 12; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSAVAA 208
 DB 149 CGSCWAFSAVAA 160

RESULT 11
 AAY49928

ID AAY49928 standard; Protein: 344 AA.

AC AAY49928;

DT 31-JAN-2000 (first entry)

DE Brassica campestris EA9 protein.

XX Brassica: regulation; development; identification; screening;

KM light promoter; seed promoter; fruit promoter; regulatory region;

KM napin; soybean; rapeseed; cotton; safflower; sunflower; expression;

KM phenotype; modification.

OS Brassica campestris.

PN US5981839-A.

XX 09-NOV-1999.

PF 07-MAR-1997; 97US-0812665.

XX 17-JAN-1985; 85US-0692605.

PP 25-JAN-1988; 88US-0147781.

PR 29-APR-1988; 88US-0188351.

PP 02-NOV-1988; 88US-0267845.

PR 09-JUN-1990; 90US-0508004.

PR 10-AUG-1993; 93US-0105852.

PR 07-JUN-1995; 95US-0484941.

PR 31-JUL-1986; 86US-0891529.

PR 26-MAY-1987; 87US-0054369.

PR 28-JUL-1987; 87US-0078538.

PR 15-MAR-1988; 88US-0168190.

PR 21-MAY-1999; 90US-0526123.

PR 14-SEP-1990; 90US-0582241.

PR 08-AUG-1991; 91US-0742834.

XX (CALJ) CALGENE LLC.

PA (CALJ) Knauf VC.

PI Knauf VC, Knauf VC.

XX WPI: 2000-012275/01.

DR N-PSDB; AAZ35693.

XX New genetic constructs useful for transforming plants with a DNA

PT sequence of interest

XX Example 5; Fig 4; 65pp; English.

PS The present invention describes genetic constructs (1) for transforming

CC plants with a DNA sequence of interest, comprising a promoter from a

CC napin, EA9 or acyl carrier protein gene, a DNA sequence of interest and

CC flanking T-DNA. The constructs can be used for the expression of

CC heterologous genes in plants. They can be used for modifying the

CC phenotype of plants. The DNA sequence of interest may encode an enzyme

CC or may comprise an antisense sequence. The constructs can be used to

CC plants such as soybean, rapeseed, cotton, safflower or sunflower. The

CC present sequence represents the Brassica campestris EA9 protein sequence.

XX Sequence 344 AA;

SO Query Match 3.0%; Score 12; DR 21; Length 344;

Best Local Similarity 100.0%; Pred. No. 0.0019;
 Matches 12; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSAVAA 208
 DB 149 CGSCWAFSAVAA 160

RESULT 12
 AAB85933

ID AAB85933 standard; Protein: 444 AA.

AC AAB85933;

DT 30-NOV-2001 (first entry)

DE B. campestris seed specific clone EA9 ORF sequence.

XX Napin gene; promoter; Brassica; phenotype; transcription; EA9.

OS Brassica campestris.

PN US6281410-B1.

XX 28-AUG-2001.

PF 15-JAN-1999; 99US-0212861.

XX 29-APR-1988; 88US-0188361.

PR 02-NOV-1988; 88US-0267845.

PR 10-AUG-1993; 93US-0105852.

PR 07-JUN-1995; 95US-0484941.

PR 31-JUL-1986; 86US-0891529.

PR 26-MAY-1987; 87US-0054369.

PR 28-JUL-1987; 87US-0078538.

PR 15-JAN-1988; 88US-0147781.

PR 21-MAR-1988; 88US-0168190.

XX (CALJ) CALGENE LLC.

PA Knauf VC, Knauf VC.

PI Knauf VC, Knauf VC.

XX WPI: 2001-54454/53.

DR N-PSDB; AAB47380.

XX Obtaining a plant that produces a seed with a modified phenotype of

PT altering a seed phenotype, comprises transforming a plant cell with a

PI DNA construct consisting of operably linked components in the direction

XX of transcription

XX Example 5; Fig 4A-E; 66pp; English.

PS The invention provides a method for obtaining a plant which produces

CC at least one seed having a modified phenotype. The method involves

CC transforming a host plant cell with a DNA construct which consists of

CC operably linked components in the direction of transcription, a promoter

CC region from a brassica plant gene, a DNA sequence of interest, either from

CC the native coding sequence, and a transcription termination region. The

CC method is useful for obtaining plants having modified phenotype or for

CC altering the phenotype of a plant seed or tissue. The DNA constructs are

CC used in manipulating plant cells to provide for regulated transcription

CC such as light inducible transcription, in a plant tissue or plant part of

CC interest of particular shapes or plant growth or in response to external

CC control. These constructs are also used for modulation of expression of

CC endogenous products as well as production of exogenous products. In the

CC seed. The present sequence represents the amino acid sequence of

CC B. campestris seed specific EA9 ORF.

XX Sequence 444 AA;

SO Query Match 3.0%; Score 12; DR 22; Length 444;

Best Local Similarity 100.0%; Pred. No. 0.0019;


```

XX 27-MAR-1997: 970S-0827171.
XX
XX
XX 27-MAR-1996: 940S-0014733
XX
XX (REGC ) UNIV CALIFORNIA
XX
XX Petersen C, Huang J:
XX
XX WPI: 2001-388923/41.
XX
XX New vaccines and cryptopain protein for diagnosis and treatment of
XX Cryptosporidium species infection
XX
XX Example 2: Fig 4; 32pp; English.
XX
XX The invention relates to vaccines and cryptopain protein for diagnosis
XX and treatment of Cryptosporidium species infection. Cryptopain protein
XX can be expressed by standard recombinant methodology. The vaccines
XX comprising the cryptopain protein or its fragments are useful for active
XX immunization of animals and humans against Cryptosporidium infection, or
XX for production of passive immune products in admixture with an adjuvant.
XX Cryptopain can be used for prophylactic, therapeutic, diagnostic and
XX detection purposes. The present sequence represents a cysteine proteinase
XX from catfish, used in hematology studies with the C. parvum cryptopain.
XX
XX Sequence 345 AA:
XX
XX Query Match 2.5%; Score 19; DB 20; Length 345.
XX Best Local Similarity 100.0%; Pred. No. 0.19;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 197 GSCWAFSAV 206
XX
XX DB 155 GSCWAFSAV 164
XX
XX
XX RESULT 16
XX AAY43540
XX ID AAY43540 standard; Protein; 443 AA.
XX
XX AC AAY43540;
XX
XX DT 26-JAN-2000 (first entry)
XX
XX DE A Leishmanian infantum cathepsin L-like cysteine proteinase, CYP.
XX
XX KW CPA; cathepsin L; cysteine proteinase; parasitic protozoa; CYP; vaccine;
XX Leishmaniasis.
XX
XX OS Leishmania infantum.
XX
XX PN W09951264-A2.
XX
XX PD 14-OCT-1999.
XX
XX PE 06-APP-1999: 95W0-CH00889.
XX
XX PR 03-APP-1998: 98GR-0007293
XX
XX PA (UNIV ) UNIV GLASGOW.
XX
XX P1 Mottram JC, Coombs GH:
XX
XX WPI: 1999-610947/52.
XX
XX DR N-PSDB; AAZ30068.
XX
XX New Leishmania vaccine comprising Leishmania proteins, useful for
XX prevention and treatment of Leishmania infection -
XX
XX Claim 19; Fig 8b; 56pp; English.
XX
XX The present sequence represents a Leishmania infantum CYP protein. CYP

```

```

XX is the major cathepsin L-like cysteine proteinase of amastigotes.
XX Proteinases have been shown to play an important role in the
XX pathogenicity of parasitic protozoa. Many Leishmania proteinases are
XX site specific. Leishmania cysteine proteinases (both active and
XX inactive), especially L. infantum CYP or CYP, are useful in the
XX preparation of vaccines for immunizing and treating mammals (especially
XX dogs and humans) against Leishmania and the disease it causes.
XX Leishmaniasis.
XX
XX Sequence 443 AA:
XX
XX Query Match 2.5%; Score 19; DB 20; Length 443.
XX Best Local Similarity 100.0%; Pred. No. 0.24;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 197 GSCWAFSAV 206
XX
XX DB 147 GSCWAFSAV 156
XX
XX
XX RESULT 17
XX AAW89559
XX ID AAW89559 standard; Protein; 472 AA.
XX
XX AC AAW89559;
XX
XX DT 16-MAR-1999 (first entry)
XX
XX DE Trifluom sp. cysteine proteinase 44.
XX
XX KW Trifluom sp. cysteine proteinase; 44.000.14614.
XX
XX OS Trifluom sp.
XX
XX PN JPI0427886-A.
XX
XX PD 15-DEC-1998.
XX
XX PE 27-MAR-1998: 98AP-0098140.
XX
XX PR 41-MAR-1997: 97AP-0114946.
XX
XX PA (SIOS ) SHOMA SANCYO CO.
XX
XX WPI: 1998-109256/10.
XX
XX DR N-PSDB; AAW82459.
XX
XX PT New DNA coding cysteine proteinase of Japanese rice wheat seed
XX useful for improving gluten for use in bakery process
XX
XX PS Claim 4; Page 20-22; 29pp; Japanese.
XX
XX CC the present sequence represents a cysteine proteinase isolated from
XX wheat seed (Trifluom sp.). The cysteine proteinase is useful for
XX improving gluten for use in the bakery process.
XX
XX Sequence 472 AA:
XX
XX Query Match 2.5%; Score 19; DB 20; Length 472.
XX Best Local Similarity 100.0%; Pred. No. 0.25;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 197 GSCWAFSAV 206
XX
XX DB 168 GSCWAFSAV 177
XX
XX
XX RESULT 18
XX AAR44529
XX ID AAR44529 standard; peptide; 14 AA.
XX
XX AC AAR44529;
XX

```


PS Claim 22: Pages 107-108; 142pp; English.

XX CC The present invention relates to coding sequences (see AAF44740-P44840
 CC and AAF44843-P44844) and proteins (see AAB65714-P65814) involved in
 CC programmed cell death (PCD; apoptosis). The coding sequences and proteins
 CC of the present invention are useful for modulating a PCD or cell death
 CC pathway and various developmental pathways in a forestry plant, by
 CC stably incorporating one of the present coding sequences into the genome
 CC of the forestry plant, where the coding sequence provides a PCD pathway
 CC that is not present in a native form of the forestry plant.

SO Sequence 200 AA:

Query Match 2.24; Score 9; Pr 22; Length 204;
 Best Local Similarity 100.0%; Prod. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 197 CGSCWAFSA 205
 |||||
 5 CGSCWAFSA 13

RESULT 21
 AAB65762
 ID AAB65762 standard; Protein: 204 AA.
 XX
 AC AAB65762;
 XX
 DI 27-MAR-2001 (first entry)
 XX
 DE Cysteine protease #3.
 XX
 KW Cell death modulator; programmed cell death; PCD; apoptosis;
 XX forestry plant.
 XX
 OS Eucalyptus grandis.
 XX
 PN M0200075331-A1.
 XX
 PD 14-DEC-2000.
 XX
 PE 02-JUN-2000; 2050W;-R200086.
 XX
 PR 04-JUN-1999; 990S-0325932.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Flinn B, Lasham A;
 XX
 DR WP1: 2001-061724/07
 DR N-PSDB; AAF44789.
 XX
 PT Novel defender against cell death polynucleotide useful for modulating
 PT programmed cell death pathway and specific development pathways in
 PT forestry plant.
 XX
 PS Claim 22: Page 104; 142pp; English.

XX CC The present invention relates to coding sequences (see AAF44740-P44840
 CC and AAF44843-P44844) and proteins (see AAB65714-P65814) involved in
 CC programmed cell death (PCD; apoptosis). The coding sequences and proteins
 CC of the present invention are useful for modulating a PCD or cell death
 CC pathway and various developmental pathways in a forestry plant, by
 CC stably incorporating one of the present coding sequences into the genome
 CC of the forestry plant, where the coding sequence provides a PCD pathway
 CC that is not present in a native form of the forestry plant.

SO Sequence 204 AA:

Query Match 2.24; Score 9; Pr 22; Length 204;
 Best Local Similarity 100.0%; Prod. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 YWLVNSW; 465
 |||||
 DB 148 YWLVNSW; 146

RESULT 22
 AAB65773
 ID AAB65773 standard; Protein: 272 AA.
 XX
 AC AAB65773;
 XX
 DI 27 MAR 2001 (first entry)
 XX
 DE Cysteine protease #14
 XX
 KW Cell death modulator; programmed cell death; PCD; apoptosis;
 KW forestry plant.
 XX
 OS Eucalyptus grandis.
 XX
 PN M0200075331-A1.
 XX
 PD 14-DEC-2000.
 XX
 PE 02-JUN-2000; 2050W;-R200086.
 XX
 PR 04-JUN-1999; 990S-0325932.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Flinn B, Lasham A;
 XX
 DR WP1: 2001-061724/07.
 DR N-PSDB; AAF44800.
 XX
 PT Novel defender against cell death polynucleotide useful for modulation
 PT programmed cell death pathway and specific development pathways in
 PT forestry plant.
 XX
 PS Claim 22: Pages 111-112; 142pp; English.

XX CC The present invention relates to coding sequences (see AAF44740-P44840
 CC and AAF44843-P44844) and proteins (see AAB65714-P65814) involved in
 CC programmed cell death (PCD; apoptosis). The coding sequences and proteins
 CC of the present invention are useful for modulating a PCD or cell death
 CC pathway and various developmental pathways in a forestry plant, by
 CC stably incorporating one of the present coding sequences into the genome
 CC of the forestry plant, where the coding sequence provides a PCD pathway
 CC that is not present in a native form of the forestry plant.

SO Sequence 272 AA:

Query Match 2.24; Score 9; Pr 22; Length 272;
 Best Local Similarity 100.0%; Prod. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 WAFSAVA 208
 |||||
 DB 16 WAFSAVA; 173

RESULT 24
 AAG36332
 ID AAG36332 standard; Protein: 213 AA.
 XX
 AC AAG36332;
 XX
 DI 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment; SGP; ID: 201-44510.


```

PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154774.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157866.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0159232.
PR 12-OCT-1999; 990S-0158369.
PR 14-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 14-OCT-1999; 990S-0159296.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161459.
PR 26-OCT-1999; 990S-0161460.
PR 26-OCT-1999; 990S-0161461.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

```

```

Query Match 2.2%; Score 9; DP 21; Length 313.
Best local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CY 200 CMASSAVAA 208
DB 121 CMASSAVAA 129

```

RESULT 24

```
AA007185
ID AA007185 standard: Protein; 425 AA.

```

```
AC AA007185;

```

```
PE 17-OCT-2000 (first entry)

```

```
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4342.

```

```

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

```

```
XX Arabidopsis thaliana.

```

```
XX OS Ep1033405-A2.

```

```

XX 06-SEP-2000.
XX 25-SEP-2000; 2000090-0301439.
XX 25-SEP-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0123180.
XX 09-MAR-1999; 990S-0123548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128214.
XX 08-APR-1999; 990S-0128714.
XX 15-APR-1999; 990S-0129245.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130349.
XX 23-APR-1999; 990S-0130510.
XX 24-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 04-MAY-1999; 990S-0132249.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 06-MAY-1999; 990S-0132487.
XX 07-MAY-1999; 990S-0132488.
XX 11-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134218.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 14-MAY-1999; 990S-0134370.
XX 18-MAY-1999; 990S-0134768.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 24-MAY-1999; 990S-0135359.
XX 25-MAY-1999; 990S-0136021.
XX 27-MAY-1999; 990S-0136392.
XX 28-MAY-1999; 990S-0136782.
XX 01-JUN-1999; 990S-0137232.
XX 03-JUN-1999; 990S-0137528.
XX 04-JUN-1999; 990S-0137702.
XX 07-JUN-1999; 990S-0137724.
XX 08-JUN-1999; 990S-0138094.
XX 10-JUN-1999; 990S-0138540.
XX 10-JUN-1999; 990S-0138647.
XX 14-JUN-1999; 990S-0139452.
XX 16-JUN-1999; 990S-0139453.
XX 16-JUN-1999; 990S-0139454.
XX 17-JUN-1999; 990S-0139455.
XX 18-JUN-1999; 990S-0139456.
XX 18-JUN-1999; 990S-0139457.
XX 18-JUN-1999; 990S-0139458.
XX 18-JUN-1999; 990S-0139459.
XX 18-JUN-1999; 990S-0139460.
XX 18-JUN-1999; 990S-0139461.
XX 18-JUN-1999; 990S-0139462.
XX 18-JUN-1999; 990S-0139463.
XX 18-JUN-1999; 990S-0139464.
XX 18-JUN-1999; 990S-0139465.
XX 21-JUN-1999; 990S-0139467.
XX 22-JUN-1999; 990S-0139899.
XX 23-JUN-1999; 990S-0140353.
XX 23-JUN-1999; 990S-0140354.
XX 24-JUN-1999; 990S-0140355.
XX 28-JUN-1999; 990S-0140823.
XX 29-JUN-1999; 990S-0140991.
XX 30-JUN-1999; 990S-0141287.
XX 01-JUL-1999; 990S-0141842.
XX 01-JUL-1999; 990S-0142154.

```


199	0.2	0.01	1.9997	9.995	0.44	2.99
198	0.2	0.01	1.9995	9.995	0.44	2.99
197	0.8	0.01	1.9995	9.995	0.44	2.94
196	0.0	0.01	1.9997	9.995	0.44	2.94
195	0.0	0.01	1.9997	9.995	0.44	2.94
194	1.2	0.01	1.9997	9.995	0.44	2.97
193	1.4	0.01	1.9997	9.995	0.44	2.94
192	1.4	0.01	1.9997	9.995	0.44	2.94
191	1.5	0.01	1.9995	9.995	0.44	2.95
190	1.6	0.01	1.9995	9.995	0.44	2.95
189	1.6	0.01	1.9995	9.995	0.44	2.95
188	1.6	0.01	1.9995	9.995	0.44	2.95
187	1.6	0.01	1.9995	9.995	0.44	2.95
186	1.6	0.01	1.9995	9.995	0.44	2.95
185	1.6	0.01	1.9995	9.995	0.44	2.95
184	1.6	0.01	1.9995	9.995	0.44	2.95
183	1.6	0.01	1.9995	9.995	0.44	2.95
182	1.6	0.01	1.9995	9.995	0.44	2.95
181	1.6	0.01	1.9995	9.995	0.44	2.95
180	1.6	0.01	1.9995	9.995	0.44	2.95
179	1.6	0.01	1.9995	9.995	0.44	2.95
178	1.6	0.01	1.9995	9.995	0.44	2.95
177	1.6	0.01	1.9995	9.995	0.44	2.95
176	1.6	0.01	1.9995	9.995	0.44	2.95
175	1.6	0.01	1.9995	9.995	0.44	2.95
174	1.6	0.01	1.9995	9.995	0.44	2.95
173	1.6	0.01	1.9995	9.995	0.44	2.95
172	1.6	0.01	1.9995	9.995	0.44	2.95
171	1.6	0.01	1.9995	9.995	0.44	2.95
170	1.6	0.01	1.9995	9.995	0.44	2.95
169	1.6	0.01	1.9995	9.995	0.44	2.95
168	1.6	0.01	1.9995	9.995	0.44	2.95
167	1.6	0.01	1.9995	9.995	0.44	2.95
166	1.6	0.01	1.9995	9.995	0.44	2.95
165	1.6	0.01	1.9995	9.995	0.44	2.95
164	1.6	0.01	1.9995	9.995	0.44	2.95
163	1.6	0.01	1.9995	9.995	0.44	2.95
162	1.6	0.01	1.9995	9.995	0.44	2.95
161	1.6	0.01	1.9995	9.995	0.44	2.95
160	1.6	0.01	1.9995	9.995	0.44	2.95
159	1.6	0.01	1.9995	9.995	0.44	2.95
158	1.6	0.01	1.9995	9.995	0.44	2.95
157	1.6	0.01	1.9995	9.995	0.44	2.95
156	1.6	0.01	1.9995	9.995	0.44	2.95
155	1.6	0.01	1.9995	9.995	0.44	2.95
154	1.6	0.01	1.9995	9.995	0.44	2.95
153	1.6	0.01	1.9995	9.995	0.44	2.95
152	1.6	0.01	1.9995	9.995	0.44	2.95
151	1.6	0.01	1.9995	9.995	0.44	2.95
150	1.6	0.01	1.9995	9.995	0.44	2.95

106	25	2003	9008	0156636																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
-----	----	------	------	---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Pattern Q-X(3)-[5E]-X-CW-X(2)-[5TAC]*

PT 009417820-A.
 XX 18-AUG-1994.
 PD
 XX 04-FEB-1994: 94WO-A000051.
 XX 05-FEB-1994: 93AT-0007109.
 XX (DARA-) DARATECH PTY LTD.
 XX
 XX Miner AR, Panaccio M, Spithill TW, Mitchell GL:
 DR WPI: 1994-279387/34.
 DR N-PSDB: AA070268.
 XX
 PT Isolated polypeptide from *Pastorales* species: useful as vaccines
 PT against a spread of infection of liver flukes
 PS Claim 14: Page 49-50; 102pp; English.
 XX
 CC An adult *F. hepatica* expression cDNA library was screened with
 CC ovine anti-cathepsin protease serum. Nine randomly selected
 CC positive clones were purified and their cDNA inserts were partially
 CC sequenced. The insert of the largest clone pFcat1 (AA070268) contains
 CC an open reading frame which is predicted to encode a protein of 326
 CC amino acids with 44% homology to human procathepsin L. The
 CC positions of pro- and pro- regions were assigned by homology to the
 CC mammalian thiol cathepsins. Recombinant pFcat1 was used to
 CC vaccinate sheep against *F. hepatica*.
 SO Sequence 326 AA;
 Query Match 2.2%; Score 9; DB 15; Length 326;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 NCSTWAFS 204
 DB 128 NCSTWAFS 136

RESULT 26
 AAP82094
 ID AAP82094 standard; protein: 333 AA.
 XX
 AC AAP82094:
 XX
 DT 17-DEC-2001 (updated)
 DT 26-OCT-1990 (first entry)
 XX
 DE pHu-16 sequence encoded human procathepsin L.
 XX
 KW human procathepsin L; pHu-16.
 XX
 OS Homo sapiens.
 XX
 PN USN7154692-N.
 XX
 PD 11-FEB-1988.
 XX
 PF 11-FEB-1988: 88US-0154692.
 XX
 PR 11-FEB-1988: 88US-0154692.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN.
 XX
 PI Gottesman MM, Gal S, Smith S;
 DR WPI: 1988-300018/42.
 DR N-PSDB: AAN82032.
 XX

Client deoxyribonucleic acid for human procathepsin L
 PT used for producing protein and antibodies and for detection of
 PT tumour activity
 XX
 XX
 PS Disclosure: 1 pp; English.
 XX
 CC This deduced protein sequence differs significantly from that for
 CC mouse procathepsin L. It is for this reason that a mouse clone
 CC cannot be used for identification or diagnostic purposes in humans.
 CC Only human DNA is appropriate for such purposes.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government owned NTS applications to prevent clashes with non-NTS
 CC granted patent numbers. For further information please visit the inventor
 CC web site at www.deptec.com (us.us.html)

SO Sequence 333 AA;
 Query Match 2.2%; Score 9; DB 9; Length 333;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 NCSTWAFS 205
 DB 135 NCSTWAFS 143

RESULT 27
 AAP28829
 ID AAP28829 standard; protein: 333 AA.
 XX
 AC AAP28829:
 XX
 DT 17-MAY-1994 (first entry)
 XX
 DE Human procathepsin L.
 XX
 KW Monoclonal antibody; cathepsin; diagnosis; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W05219756-A.
 XX
 PD 12-NOV-1992.
 XX
 PE 30-APR-1992: 92WO-0000568.
 XX
 PR 02-MAY-1991: 91JP-0128274.
 XX
 PA (IDPK) IDEMITSU KOSAN CO. LTD.
 XX
 PI Irie S, Kasahara T, Miyajima T, Nakayama K, Yokoi H;
 DR WPI: 1992-398880/48.
 DR N-PSDB: AA031260.
 XX
 PT Monoclonal antibodies specific for human procathepsin L and
 PT five peptides derived from it, useful in diagnosis and
 PT treatment of liver, lung, uterine and ovarian cancers
 XX
 PS Claim 1: Page 35; 48pp; Japanese.
 XX

Proctor and his colleagues have published their findings in the *Journal of Interpersonal Violence*. They also used the findings to develop a guide for therapists.

ANAL. Calcd for $C_{10}H_{10}O$: C, 88.10%; H, 7.39%. Found: C, 88.1%; H, 7.4%.

$$\frac{d}{dt} \left(\int_{\Omega} M_0(t) dx \right) = - \int_{\Omega} M_0(t) dx + \int_{\partial \Omega} M_0(t) dx$$

Author	Year	Sample Size	Method	Findings
Smith et al.	2015	1,200	Survey	High levels of stress and anxiety reported among students.
Johnson et al.	2016	800	Interview	Students perceive a lack of support from faculty and staff.
Williams et al.	2017	1,500	Survey	Increased mental health issues since the start of the semester.
Miller et al.	2018	900	Focus Group	Students feel overwhelmed by the workload and deadlines.
Davis et al.	2019	1,100	Survey	Significant improvement in student satisfaction after intervention.
Garcia et al.	2020	1,300	Survey	Continued challenges with remote learning and social isolation.
Lee et al.	2021	1,400	Survey	Positive impact of peer support programs on student well-being.
White et al.	2022	1,600	Survey	Need for more comprehensive mental health resources.
Black et al.	2023	1,700	Survey	Successful implementation of new campus safety protocols.
Green et al.	2024	1,800	Survey	Enhanced student engagement through digital platforms.
Brown et al.	2025	1,900	Survey	Improved academic performance following targeted interventions.

[illegible][illegible]

\mathbb{R}^n

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

$$X^{\alpha} = (X^{\alpha_1}, \dots, X^{\alpha_n})^T, \quad Y^{\alpha} = (Y^{\alpha_1}, \dots, Y^{\alpha_n})^T,$$
[illegible]

ΔA
 ΔB
 ΔC
 ΔD
 ΔE
 ΔF
 ΔG
 ΔH
 ΔI
 ΔJ
 ΔK
 ΔL
 ΔM
 ΔN
 ΔO
 ΔP
 ΔQ
 ΔR
 ΔS
 ΔT
 ΔU
 ΔV
 ΔW
 ΔX
 ΔY
 ΔZ

$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

[illegible]

Figure	Figure
Figure 1	Figure 2
Figure 3	Figure 4
Figure 5	Figure 6
Figure 7	Figure 8
Figure 9	Figure 10
Figure 11	Figure 12
Figure 13	Figure 14
Figure 15	Figure 16
Figure 17	Figure 18
Figure 19	Figure 20
Figure 21	Figure 22
Figure 23	Figure 24
Figure 25	Figure 26
Figure 27	Figure 28
Figure 29	Figure 30
Figure 31	Figure 32
Figure 33	Figure 34
Figure 35	Figure 36
Figure 37	Figure 38
Figure 39	Figure 40
Figure 41	Figure 42
Figure 43	Figure 44
Figure 45	Figure 46
Figure 47	Figure 48
Figure 49	Figure 50
Figure 51	Figure 52
Figure 53	Figure 54
Figure 55	Figure 56
Figure 57	Figure 58
Figure 59	Figure 60
Figure 61	Figure 62
Figure 63	Figure 64
Figure 65	Figure 66
Figure 67	Figure 68
Figure 69	Figure 70
Figure 71	Figure 72
Figure 73	Figure 74
Figure 75	Figure 76
Figure 77	Figure 78
Figure 79	Figure 80
Figure 81	Figure 82
Figure 83	Figure 84
Figure 85	Figure 86
Figure 87	Figure 88
Figure 89	Figure 90
Figure 91	Figure 92
Figure 93	Figure 94
Figure 95	Figure 96
Figure 97	Figure 98
Figure 99	Figure 100
Figure 101	Figure 102
Figure 103	Figure 104
Figure 105	Figure 106
Figure 107	Figure 108
Figure 109	Figure 110
Figure 111	Figure 112
Figure 113	Figure 114
Figure 115	Figure 116
Figure 117	Figure 118
Figure 119	Figure 120
Figure 121	Figure 122
Figure 123	Figure 124
Figure 125	Figure 126
Figure 127	Figure 128
Figure 129	Figure 130
Figure 131	Figure 132
Figure 133	Figure 134
Figure 135	Figure 136
Figure 137	Figure 138
Figure 139	Figure 140
Figure 141	Figure 142
Figure 143	Figure 144
Figure 145	Figure 146
Figure 147	Figure 148
Figure 149	Figure 150
Figure 151	Figure 152
Figure 153	Figure 154
Figure 155	Figure 156
Figure 157	Figure 158
Figure 159	Figure 160
Figure 161	Figure 162
Figure 163	Figure 164
Figure 165	Figure 166
Figure 167	Figure 168
Figure 169	Figure 170
Figure 171	Figure 172
Figure 173	Figure 174
Figure 175	Figure 176
Figure 177	Figure 178
Figure 179	Figure 180
Figure 181	Figure 182
Figure 183	Figure 184
Figure 185	Figure 186
Figure 187	Figure 188
Figure 189	Figure 190
Figure 191	Figure 192
Figure 193	Figure 194
Figure 195	Figure 196
Figure 197	Figure 198
Figure 199	Figure 200
Figure 201	Figure 202
Figure 203	Figure 204
Figure 205	Figure 206
Figure 207	Figure 208
Figure 209	Figure 210
Figure 211	Figure 212
Figure 213	Figure 214
Figure 215	Figure 216
Figure 217	Figure 218
Figure 219	Figure 220
Figure 221	Figure 222
Figure 223	Figure 224
Figure 225	Figure 226
Figure 227	Figure 228
Figure 229	Figure 230
Figure 231	Figure 232
Figure 233	Figure 234
Figure 235	Figure 236
Figure 237	Figure 238
Figure 239	Figure 240
Figure 241	Figure 242
Figure 243	Figure 244
Figure 245	Figure 246
Figure 247	Figure 248
Figure 249	Figure 250
Figure 251	Figure 252
Figure 253	Figure 254
Figure 255	Figure 256
Figure 257	Figure 258
Figure 259	Figure 260
Figure 261	Figure 262
Figure 263	Figure 264
Figure 265	Figure 266
Figure 267	Figure 268
Figure 269	Figure 270
Figure 271	Figure 272
Figure 273	Figure 274
Figure 275	Figure 276
Figure 277	Figure 278
Figure 279	Figure 280
Figure 281	Figure 282
Figure 283	Figure 284
Figure 285	Figure 286
Figure 287	Figure 28

11. *Phragmites australis* (Cav.) Trin. ex Steud.

X

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100.

III. NISHIMOTO AND KATO

of DNA and the related literature. The authors are grateful to the National Natural Science Foundation of China (Grant No. 30070010) for the financial support of this work.

[illegible]

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

[illegible]
$$N_{\text{eff}} = 3.36 \pm 0.17 \pm 0.04$$

Figure 1. The effect of the concentration of the *Agrobacterium* suspension on the transformation efficiency of *Agrobacterium* strains. The *Agrobacterium* strains were grown in YEA medium for 24 h at 28°C. The cell concentration was adjusted to 1.0 × 10⁸ cells/ml. The cell suspension was mixed with the plant tissue and incubated for 24 h at 28°C. The plant tissue was then cultured on the selective medium. The transformation efficiency was determined as the number of transformants per 100 mg of plant tissue. The data are the mean ± SD of three independent experiments.

Table 1. *Continued*

1. *Phragmites australis* (Cav.) Trin. ex Steud.

[illegible]

1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 26

KW metabolic disorder; hormonal disorder; immune disorder; wound;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9918126-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 06-OCT-1998; 98WO-0904514
 XX
 PR 07-OCT-1997; 97JP-0274674.
 XX
 PA (ONKOY) ONO PHARM CO LTD.
 XX
 PI Fukushima D, Shibayama S, Tada H;
 XX
 DR WPI: 1999-272254/23.
 XX
 DR N-PSDB: AAX35694, AAX35695.
 XX
 PT Polypeptides identified by the signal sequence trap method from a
 PT human cDNA library
 PS
 PS Claim 1: Page 65-66; 281pp; Japanese
 XX
 CC AAY0358-84 represents novel polypeptides which are identified from a
 CC human placental cDNA library by the signal sequence trap (SST) method.
 CC The polypeptides are encoded by the cDNA sequences in AAX35694-X35747.
 CC The polypeptides have a broad range of physiological activity, including
 CC immunisation against and inhibition of infections, allergies and cancer;
 CC regulation of tissue formation and repair; activin/inhibin activity;
 CC chemokine/cytokine activity; blood coagulation regulation; and
 CC receptor/ligand agonist or antagonist activity. The polypeptides can
 CC be used for prevention and treatment of disorders including infections
 CC by bacteria, yeasts and viruses (including HIV) and protozoa, metabolic
 CC and hormonal disorders, immune disorders (including severe combined
 CC immunodeficiency (SCID) and AIDS; thrombosis; cancer; and traumatic or
 CC surgical wounds.
 XX
 SC Sequence 334 AA;
 Query Match 2.28; Score 9; DB 20; Length 334;
 Best Local Similarity 100.0%, Pred. No. 1 9,
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 197 CGSCWAFSA 205
 DB 135 CGSCWAFSA 143
 RESULT 31
 AAY94300
 ID AAY94300 standard; Protein: 334 AA.
 XX
 AC AAY94300;
 XX
 DT 28-APR-1999 (first entry)
 XX
 DE Human cathepsin (LCAP).
 KW Cathepsin; LCAP; celi proliferation disorder; cancer; adenocarcinoma;
 KW sarcoma; leukaemia; melanoma; myeloma; teratocarcinoma; immune disorder;
 KW lymphoma; Addison's disease; AIDS; adult respiratory distress syndrome;
 KW allergy; asthma; Crohn's disease; multiple sclerosis; arthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO9900508-A1.
 XX
 PD 07-JAN-1999.
 XX
 PF 19-JUN-1998; 98WO-0812863.
 XX

PP 26-JUN-1997; 97US-092356.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Bandman O, Corley NC, Guegler KJ, Shah P;
 XX
 DR WPI: 1999-095753/08.
 XX
 DR N-PSDB: AAX05743.
 XX
 PT Human cathepsin, LCAP - used for treating disorders associated with
 PT cell proliferation
 PS
 PS Claim 1: Fig 1A-D; 60pp; English.
 XX
 CC This represents a human cathepsin. Host cells containing an expression
 CC vector comprising the LCAP nucleic acid are used for the recombinant
 CC production of the protein. LCAP can be used to prevent or treat a
 CC disorder associated with cell proliferation. These include various types
 CC of cancer, such as adenocarcinoma, sarcoma, lymphoma, leukaemia,
 CC melanoma, myeloma, teratocarcinoma, and cancers of the adrenal gland,
 CC bladder, bone, brain, breast, gastrointestinal tract, heart, kidney,
 CC liver, lung, ovary, pancreas, parathyroid, parathyroid, prostate,
 CC salivary glands, skin, spleen, testis, thyroid and uterus. LCAP may
 CC also be used to prevent or treat an immune disorder, such as Addison's
 CC disease, AIDS, adult respiratory distress syndrome, allergy, asthma,
 CC Crohn's disease, emphysema, multiple sclerosis, scleroderma, and
 CC osteoporosis, psoriasis, rheumatoid arthritis, scleroderma, and
 CC autoimmune thyroiditis.
 XX
 SC Sequence 334 AA;
 Query Match 2.28; Score 9; DB 20; Length 334;
 Best Local Similarity 100.0%, Pred. No. 1 9,
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 197 CGSCWAFSA 205
 DB 135 CGSCWAFSA 143
 RESULT 32
 AAY81487
 ID AAY81487 standard; Protein: 334 AA.
 XX
 AC AAY81487;
 XX
 DT 18-JUN-2000 (first entry)
 XX
 DE Human cathepsin L2.
 KW Cathepsin L2; human; brain; cysteine protease; expressed sequence tag;
 KW BSH; cancer; Alzheimer's disease; emphysema; rheumatoid arthritis;
 KW muscular dystrophy; osteoporosis; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN KJ
 XX
 DT Peptide
 XX
 DE Protein
 XX
 E1 /note="Signal peptide"
 E1 /note="Mature human cathepsin L2"
 XX
 PF JF200005086-A.
 XX
 PD 22-FEB-2000.
 XX
 PF 03-JUN-1999; 99JP-0156945.
 XX
 PR 05-JUN-1998; 98JP-0172147.
 XX
 PA (FUJY) FUJII PHARM IND CO LTD.
 XX
 DR WPI: 2000-353225/31.

XX 11-OCT-2001.
PD 30-MAR-2001; 4001WO-US08631.
XX 31-MAR-2001; 2000NS-0640217
XX 23-AUG-2001; 2000NS-0643167
XX (HYSE-) HXSEQ INV
PI Dmanac RT, Liu C, Tang YT;
DR WPI: 2001-63362773
DR N-PSDB; AAS85613.
XX
XX
XX New isolated polynucleotide and encoded polypeptides useful in
PT diagnostics; forensic gene mapping; identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 20. SEQ ID NO 51785; 103pp; English
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polynucleotide (I) sequences. (II) is useful as hybridisation probes,
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore lost activity of (I) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its biologic partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostic, forensic, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on ENA and
CC amino acid sequences. Arabidopsis thaliana represents novel human
CC diagnostic amino acid sequences of the invention
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WPI
CC at ftp.wpi.edu/pub/published_pat_sequences
XX
SQ Sequence 336 AA;
Query Match 2.28; Score 9; DB 23; Length 336;
Best Local Similarity 100.0%; Prod No 1 9;
Matches 9; Conserved 0; Mismatches 0; Gaps 0;
OY 197 CGSCWAFSA 205
DB 137 CGSCWAFSA 145
RESULT 35
AAG36331
ID AAG36331 standard; protein; 341 AA.
XX
XX AAG36331;
AC
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO. 44509
DE
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX

PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0122180.
XX 09-MAR-1999; 990S-0122518.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126261.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129485.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0130510.
XX 24-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 30-APR-1999; 990S-0132247.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 07-MAY-1999; 990S-0132487.
XX 11-MAY-1999; 990S-0132953.
XX 14-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134218.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 14-MAY-1999; 990S-0134270.
XX 18-MAY-1999; 990S-0134768.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 24-MAY-1999; 990S-0135629.
XX 25-MAY-1999; 990S-0135831.
XX 27-MAY-1999; 990S-0136292.
XX 28-MAY-1999; 990S-0136782.
XX 01-JUN-1999; 990S-0137252.
XX 03-JUN-1999; 990S-0137528.
XX 04-JUN-1999; 990S-0137602.
XX 07-JUN-1999; 990S-0137724.
XX 08-JUN-1999; 990S-0138094.
XX 10-JUN-1999; 990S-0138540.
XX 10-JUN-1999; 990S-0138847.
XX 14-JUN-1999; 990S-0139119.
XX 16-JUN-1999; 990S-0139452.
XX 17-JUN-1999; 990S-0139453.
XX 18-JUN-1999; 990S-0139454.
XX 18-JUN-1999; 990S-0139455.
XX 18-JUN-1999; 990S-0139456.
XX 18-JUN-1999; 990S-0139457.
XX 18-JUN-1999; 990S-0139458.
XX 18-JUN-1999; 990S-0139459.
XX 18-JUN-1999; 990S-0139460.
XX 18-JUN-1999; 990S-0139461.
XX 18-JUN-1999; 990S-0139462.
XX 18-JUN-1999; 990S-0139463.
XX 18-JUN-1999; 990S-0139464.
XX 18-JUN-1999; 990S-0139465.
XX 19-JUN-1999; 990S-0139466.
XX 21-JUN-1999; 990S-0139467.
XX 22-JUN-1999; 990S-0139468.
XX 23-JUN-1999; 990S-0140353.
XX 23-JUN-1999; 990S-0140354.
XX 24-JUN-1999; 990S-0140645.
XX 28-JUN-1999; 990S-0140823.
XX 29-JUN-1999; 990S-0140991.
XX 30-JUN-1999; 990S-0141287.
XX 01-JUL-1999; 990S-0141842.
XX 02-JUL-1999; 990S-0142154.
XX 02-JUL-1999; 990S-0142055.

AAV43538
ID AAV43538 standard; Protein; 354 AA.
XX
AC AAV43538;
XX
DT 26-JAN-2000 (first entry)
XX
DE A Leishmania infantum cathepsin L-like cysteine proteinase, CPA.
XX
KW CPA; cathepsin L; cysteine proteinase; furin-like protease; CPB; vacuole;
XX
OS Leishmania infantum.
XX
FN W09951264-A2.
XX
PD 14-OCT-1999.
XX
PE 06-APR-1999; 99MO-G800889.
XX
PF 01-APR-1998; 98GB-0007294.
XX
PA (UNIV) UNIV GLASGOW.
XX
PI Moltram JC, Coombs GH;
XX
DR WP1: 1999-610947/52.
XX
DE N-PSDB; AAZ30066.
XX
PT New Leishmania vaccine comprising Leishmania protein; useful for
XX
PT prevention and treatment of Leishmania infection.
XX
PS Claim 19; Fig 6; 56pp; English.
XX
CC The present sequence represents a Leishmania infantum CPA protein. CPA
XX
CC is a non-abundant cathepsin L-like cysteine proteinase. Proteinases
XX
CC have been shown to play an important role in the pathogenicity of
XX
CC parasitic protozoa. Many Leishmania proteinases are stage specific.
XX
CC Leishmania cysteine proteinases (both active and inactive), especially
XX
CC L. infantum CPA or CPB, are useful in the preparation of vaccines for
XX
CC immunising and treating mammals (especially dogs and humans) against
XX
CC Leishmania and the disease it causes, leishmaniasis.
XX
SO Sequence 354 AA;
XX
Query Match 2.2%; Score 9; DB 20; Length 354;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 CGSCWAFSA 205
DB 150 CGSCWAFSA 158

RESULT 42
AAV33689
ID AAV33689 standard; Protein; 355 AA.
XX
AC AAV33689;
XX
DT 12-JAN-2000 (first entry)
XX
DE Maize 34Pda cysteine proteinase M1rl protein fragment.
XX
KW Cysteine proteinase; M1rl; Maize; sterility; male plant; anther; cotton;
XX
KW insect resistance; lepidoptera; sugar cane; sugar beet; seed.
XX
OS Zea mays.
XX
OS US5977440-A.
XX
PN 02-NOV-1999.
XX

PF 30-APR-1996; 96DS-064114.
XX
XX
PR 30-APR-1996; 96DS-064114.
XX
PA (JOHN/) JUNE H S.
XX
FA (WILL/) WILLIAMS W P.
XX
PA (JIAN/) JIAN H.
XX
PA (PECH/) PECHAN T.
XX
PI Williams WP, Junge DS, Jian H, Pechan T;
XX
DR WP1: 1999-619723/53.
XX
DE N-PSDB; AAZ4708.
XX
PT cDNA encoding maize cysteine protease useful for imparting insect
XX
PT resistance or male sterility to plants, especially maize.
XX
PS Disclosure; Fig 1A-C; 12pp; English.
XX
XX
CC This invention describes a novel cDNA, called M1rl, encoding a maize
XX
CC 33 kD cysteine proteinase. An expression vector containing the entire
XX
CC M1rl sequence linked to one or more control elements can be used to
XX
CC impart sterility to a male plant by a method comprising transforming an
XX
CC anther of the plant with the vector and expressing the 33 kD cysteine
XX
CC proteinase in the anther. A bacteriophage transformed with such a vector
XX
CC can be used to impart resistance to a plant by a method comprising
XX
CC infecting the plant with the transformed bacteriophage. A seed containing
XX
CC a RNA molecule consisting of nucleotides 987-1250 of the M1rl sequence
XX
CC linked to a cDNA molecule encoding a cysteine proteinase and one or more
XX
CC control elements can be used to impart insect resistance to a plant.
XX
CC susceptible to lepidopteran feeding, preferably a cotton, sugarcane,
XX
CC sugar beet or especially maize plant, by a method comprising introducing
XX
CC the vector into a plant seed and germinating the seed into a plant. This
XX
CC sequence represents the maize M1rl protein described in the method of the
XX
XX
SO Sequence 355 AA;
XX
Query Match 2.2%; Score 9; DB 20; Length 355;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 200 CWAFSAVAA 208
DB 138 CWAFSAVAA 146

RESULT 43
AAB65760
ID AAB65760 standard; Protein; 386 AA.
XX
AC AAB65760;
XX
DT 27-MAR-2001 (first entry)
XX
DE Cysteine protease #1.
XX
KW Cell death modulator; programmed cell death; PCP; apoptosis;
XX
KW forestry plant.
XX
OS Eucalyptus grandis.
XX
PD W020007331-A1.
XX
PD 14-DIC-2000.
XX
PF 02-JUN-2000; 2000MO-N200086.
XX
PR 04-JUN-1999; 99US-0325932.
XX
PA (GENE) GENESIS RES & DEV CORP LTD,
XX
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX

PR 22-JUL-1999; 9905-0145192
 PR 23-JUL-1999; 9905-0145146
 PR 23-JUL-1999; 9905-0145218
 PR 24-JUL-1999; 9905-0145224
 PR 26-JUL-1999; 9905-0145276
 PR 27-JUL-1999; 9905-0145313
 PR 27-JUL-1999; 9905-0145318
 PR 27-JUL-1999; 9905-0145319
 PR 28-JUL-1999; 9905-0145351
 PR 02-AUG-1999; 9905-0146386
 PR 02-AUG-1999; 9905-0146388
 PR 02-AUG-1999; 9905-0146389
 PR 03-AUG-1999; 9905-0147038
 PR 04-AUG-1999; 9905-0147204
 PR 04-AUG-1999; 9905-0147302
 PR 05-AUG-1999; 9905-0147192
 PR 06-AUG-1999; 9905-0147303
 PR 06-AUG-1999; 9905-0147416
 PR 09-AUG-1999; 9905-0147404
 PR 09-AUG-1999; 9905-0147935
 PR 10-AUG-1999; 9905-0148171
 PR 11-AUG-1999; 9905-0148319
 PR 12-AUG-1999; 9905-0148341
 PR 13-AUG-1999; 9905-0148565
 PR 13-AUG-1999; 9905-0148684
 PR 16-AUG-1999; 9905-0148688
 PR 17-AUG-1999; 9905-0149175
 PR 18-AUG-1999; 9905-0149476
 PR 20-AUG-1999; 9905-0149723
 PR 20-AUG-1999; 9905-0149929
 PR 23-AUG-1999; 9905-0149902
 PR 23-AUG-1999; 9905-0149930
 PR 25-AUG-1999; 9905-0150556
 PR 26-AUG-1999; 9905-0150884
 PR 27-AUG-1999; 9905-0151065
 PR 27-AUG-1999; 9905-0151066
 PR 27-AUG-1999; 9905-0151080
 PR 30-AUG-1999; 9905-0151303
 PR 31-AUG-1999; 9905-0151418
 PR 01-SEP-1999; 9905-0151930
 PR 07-SEP-1999; 9905-0152463
 PR 10-SEP-1999; 9905-0153070
 PR 13-SEP-1999; 9905-0153258
 PR 15-SEP-1999; 9905-0154018
 PR 16-SEP-1999; 9905-0154039
 PR 20-SEP-1999; 9905-0154774
 PR 22-SEP-1999; 9905-0155139
 PR 23-SEP-1999; 9905-0155486
 PR 24-SEP-1999; 9905-0155679
 PR 28-SEP-1999; 9905-0156458
 PR 29-SEP-1999; 9905-0156596
 PR 04-OCT-1999; 9905-0157117
 PR 05-OCT-1999; 9905-0157253
 PR 06-OCT-1999; 9905-0157865
 PR 07-OCT-1999; 9905-0158029
 PR 08-OCT-1999; 9905-0158233
 PR 12-OCT-1999; 9905-0158466
 PR 13-OCT-1999; 9905-0158294
 PR 13-OCT-1999; 9905-0159294
 PR 13-OCT-1999; 9905-0159295
 PR 14-OCT-1999; 9905-0159329
 PR 14-OCT-1999; 9905-0159330
 PR 14-OCT-1999; 9905-0159331
 PR 14-OCT-1999; 9905-0159637
 PR 14-OCT-1999; 9905-0159638
 PR 18-OCT-1999; 9905-0159584
 PR 21-OCT-1999; 9905-0160741
 PR 21-OCT-1999; 9905-0160767
 PR 21-OCT-1999; 9905-0160768
 PR 21-OCT-1999; 9905-0160770
 PR 21-OCT-1999; 9905-0160814

PR 21-OCT-1999; 9905-0160815
 PR 22-OCT-1999; 9905-0160980
 PR 22-OCT-1999; 9905-0160981
 PR 22-OCT-1999; 9905-0160984
 PR 25-OCT-1999; 9905-0161404
 PR 25-OCT-1999; 9905-0161405
 PR 25-OCT-1999; 9905-0161406
 PR 26-OCT-1999; 9905-0161359
 PR 26-OCT-1999; 9905-0161360
 PR 26-OCT-1999; 9905-0161361
 PR 28-OCT-1999; 9905-0161920
 PR 28-OCT-1999; 9905-0161992
 PR 28-OCT-1999; 9905-0161994
 PR 29-OCT-1999; 9905-0162142

Query: Maf Th 2.2K; Score 92 PR 41; Length 411;
 Best local similarity: 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Totals 0; Gaps 0

QY 197 COSMAMSA 205
 DB 109 COSMAMSA 117

RESULT 45
 AAC39792
 ID AAC39792 standard; Protein: 452 AA.
 XX
 AC AAC39792;
 XX
 YZ 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SIV to No. 49288.
 XX
 KW Protein identification, signal transduction pathway; metabolic pathway;
 FM hybridisation assay; gene; maff109; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP103405-AZ.
 XX
 PU 06-SEP-2000.
 XX
 PU 25-FEB-2000; 2000BP-0401439.
 XX
 PR 25-FEB-1999; 9905-0121825
 PR 05-MAR-1999; 9905-0123180
 PR 09-MAR-1999; 9905-0124546
 PR 23-MAR-1999; 9905-0125788
 PR 25-MAR-1999; 9905-0126264
 PR 29-MAR-1999; 9905-0126785
 PR 01-APR-1999; 9905-0127462
 PR 06-APR-1999; 9905-0128234
 PR 08-APR-1999; 9905-0128714
 PR 16-APR-1999; 9905-0129845
 PR 19-APR-1999; 9905-0130927
 PR 21-APR-1999; 9905-0133449
 PR 23-APR-1999; 9905-0133510
 PR 23-APR-1999; 9905-0130841
 PR 28-APR-1999; 9905-0131449
 PR 30-APR-1999; 9905-0132048
 PR 30-APR-1999; 9905-0132407
 PR 04-MAY-1999; 9905-0132484
 PR 05-MAY-1999; 9905-0132485
 PR 06-MAY-1999; 9905-0132486
 PR 06-MAY-1999; 9905-0132487
 PR 07-MAY-1999; 9905-0132488
 PR 11-MAY-1999; 9905-0132566
 PR 14-MAY-1999; 9905-0134218
 PR 14-MAY-1999; 9905-0134219
 PR 14-MAY-1999; 9905-0134221
 PR 14-MAY-1999; 9905-0134370

[illegible]

09 0000 WAFSAVAA 000
1111111
111 111 WAFSAVAA 111

090000 000000 000000 000000 000000
000000 000000 000000 000000 000000

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:46:11 : Search time 23 seconds

(without alignments)
1675,297 Million cell updates/sec

Title: us-09-598-062-4

Perfect score: 401

Sequence: 1 MDTGNVEEHGYSYPA.....KSTGGLVEVTVVNGSI 101

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 28138 seqs, 96089434 residues

Word size : 0

Total number of hits satisfying chosen parameters: 28138

Minimum DR seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database :

1: p1r71:*
2: p1r72:*
3: p1r73:*
4: p1r74:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No | Score | Query Match | Length | DR | ID | Description |
|-----------|-------|-------------|--------|----|--------|----------------------|
| 1 | 12 | 3.0 | 446 | 2 | JAD159 | cysteine proteinase |
| 2 | 12 | 3.0 | 466 | 2 | 709416 | cysteine proteinase |
| 3 | 11 | 2.7 | 348 | 2 | JN0938 | cathepsin (EC 3.4.22 |
| 4 | 11 | 2.7 | 367 | 2 | JN0634 | cathepsin (EC 3.4.22 |
| 5 | 10 | 2.5 | 217 | 2 | S16044 | cathepsin K (EC 3. |
| 6 | 10 | 2.5 | 314 | 2 | S68782 | cathepsin L (EC 3. |
| 7 | 10 | 2.5 | 345 | 1 | PIPA | cathepsin L (EC 3. |
| 8 | 10 | 2.5 | 443 | 2 | S29245 | cathepsin L (EC 3. |
| 9 | 10 | 2.5 | 444 | 2 | A48566 | cysteine proteinase |
| 10 | 10 | 2.5 | 467 | 2 | T23720 | hypothetical prote |
| 11 | 10 | 2.5 | 471 | 1 | KHE708 | erythrin (EC 3.4.22 |
| 12 | 9 | 2.2 | 183 | 2 | A44938 | cysteine proteinase |
| 13 | 9 | 2.2 | 313 | 2 | S47433 | cathepsin L (EC 3. |
| 14 | 9 | 2.2 | 314 | 2 | S41438 | cathepsin L (EC 3. |
| 15 | 9 | 2.2 | 317 | 2 | S41451 | cysteine proteinase |
| 16 | 9 | 2.2 | 320 | 2 | S17651 | cathepsin L (EC 3. |
| 17 | 9 | 2.2 | 324 | 2 | S47432 | cathepsin L (EC 3. |
| 18 | 9 | 2.2 | 325 | 2 | S49451 | cysteine proteinase |
| 19 | 9 | 2.2 | 326 | 2 | S41991 | cathepsin L (EC 3. |
| 20 | 9 | 2.2 | 330 | 2 | A45087 | cathepsin S (EC 3. |
| 21 | 9 | 2.2 | 333 | 1 | KHNU | cathepsin S (EC 3. |
| 22 | 9 | 2.2 | 334 | 1 | KHNTL | cathepsin L (EC 3. |
| 23 | 9 | 2.2 | 334 | 1 | KHNSL | cathepsin L (EC 3. |
| 24 | 9 | 2.2 | 334 | 2 | A58195 | cathepsin L (EC 3. |
| 25 | 9 | 2.2 | 337 | 2 | T23187 | cathepsin L (EC 3. |
| 26 | 9 | 2.2 | 340 | 2 | T12516 | fruit bromelain (E |
| 27 | 9 | 2.2 | 341 | 2 | T45979 | probable cysteine |
| 28 | 9 | 2.2 | 343 | 2 | D86198 | cysteine proteinase |
| 29 | 9 | 2.2 | 345 | 2 | D84752 | probable cysteine |

| | | | | | | |
|-----|---|-----|-----|---|--------|----------------------|
| 30 | 9 | 2.2 | 345 | 2 | F21856 | probable cysteine |
| 31 | 9 | 2.2 | 348 | 2 | F84672 | probable cysteine |
| 32 | 9 | 2.2 | 349 | 2 | F24292 | cysteine proteinase |
| 33 | 9 | 2.2 | 352 | 2 | T10501 | fruit bromelain (E |
| 34 | 9 | 2.2 | 354 | 2 | S25267 | cysteine proteinase |
| 35 | 9 | 2.2 | 364 | 2 | T12039 | cysteine proteinase |
| 36 | 9 | 2.2 | 364 | 2 | T46630 | cysteine proteinase |
| 37 | 9 | 2.2 | 367 | 2 | T06529 | cysteine proteinase |
| 38 | 9 | 2.2 | 368 | 2 | S47312 | cysteine proteinase |
| 39 | 9 | 2.2 | 371 | 2 | S59597 | cysteine proteinase |
| 40 | 9 | 2.2 | 427 | 2 | S57776 | cysteine proteinase |
| 41 | 9 | 2.2 | 447 | 2 | S70248 | cysteine proteinase |
| 42 | 9 | 2.2 | 455 | 2 | T12041 | cysteine proteinase |
| 43 | 9 | 2.2 | 458 | 1 | KHRC7A | cathepsin (EC 3.4.22 |
| 44 | 9 | 2.2 | 464 | 2 | S44602 | cysteine proteinase |
| 45 | 9 | 2.2 | 467 | 2 | A60667 | cysteine proteinase |
| 46 | 9 | 2.2 | 467 | 2 | A45629 | cysteine proteinase |
| 47 | 9 | 2.2 | 493 | 2 | T01206 | cysteine proteinase |
| 48 | 9 | 2.2 | 498 | 2 | T41871 | hypothetical prote |
| 49 | 9 | 2.2 | 53 | 2 | F42946 | cysteine proteinase |
| 50 | 9 | 2.2 | 156 | 2 | T02166 | cysteine proteinase |
| 51 | 9 | 2.2 | 166 | 2 | B44948 | cysteine proteinase |
| 52 | 9 | 2.2 | 182 | 2 | D82925 | hypothetical prote |
| 53 | 9 | 2.2 | 211 | 2 | S21864 | probable cysteine |
| 54 | 9 | 2.2 | 214 | 2 | S44476 | cathepsin L (EC 3. |
| 55 | 9 | 2.2 | 218 | 1 | KH93B | cathepsin L (EC 3. |
| 56 | 9 | 2.2 | 218 | 2 | S67481 | cathepsin L (EC 3. |
| 57 | 9 | 2.2 | 221 | 2 | A59041 | cysteine proteinase |
| 58 | 9 | 2.2 | 245 | 2 | J00347 | allergen hor p 1 - |
| 59 | 9 | 2.2 | 260 | 2 | B42957 | probable food-bor |
| 60 | 9 | 2.2 | 278 | 2 | S44425 | cysteine proteinase |
| 61 | 9 | 2.2 | 308 | 2 | T46234 | hypothetical prote |
| 62 | 9 | 2.2 | 319 | 2 | A61500 | allergen hor 1 - p |
| 63 | 9 | 2.2 | 322 | 2 | S19649 | cysteine proteinase |
| 64 | 9 | 2.2 | 323 | 2 | S19650 | cysteine proteinase |
| 65 | 9 | 2.2 | 326 | 2 | S51027 | cathepsin L (EC 3. |
| 66 | 9 | 2.2 | 326 | 2 | T09259 | cathepsin L (EC 3. |
| 67 | 9 | 2.2 | 329 | 2 | J01121 | cysteine proteinase |
| 68 | 9 | 2.2 | 328 | 2 | T26445 | hypothetical prote |
| 69 | 9 | 2.2 | 329 | 2 | J07245 | cathepsin K (EC 3. |
| 70 | 9 | 2.2 | 329 | 2 | A49658 | cathepsin K (EC 3. |
| 71 | 9 | 2.2 | 329 | 2 | S24227 | cathepsin K (EC 3. |
| 72 | 9 | 2.2 | 329 | 2 | A53810 | cathepsin L (EC 3. |
| 73 | 9 | 2.2 | 342 | 2 | S71774 | cysteine proteinase |
| 74 | 9 | 2.2 | 348 | 2 | S34494 | cysteine proteinase |
| 75 | 9 | 2.2 | 352 | 2 | T07766 | cysteine proteinase |
| 76 | 9 | 2.2 | 352 | 2 | T06122 | cysteine proteinase |
| 77 | 9 | 2.2 | 357 | 2 | A86341 | cysteine proteinase |
| 78 | 9 | 2.2 | 357 | 2 | S49166 | cysteine proteinase |
| 79 | 9 | 2.2 | 358 | 2 | S42882 | cysteine proteinase |
| 80 | 9 | 2.2 | 360 | 2 | S57277 | cysteine proteinase |
| 81 | 9 | 2.2 | 360 | 2 | T08122 | cysteine proteinase |
| 82 | 9 | 2.2 | 361 | 2 | T06708 | cysteine proteinase |
| 83 | 9 | 2.2 | 362 | 2 | S12581 | cysteine proteinase |
| 84 | 9 | 2.2 | 362 | 2 | S22532 | cysteine proteinase |
| 85 | 9 | 2.2 | 362 | 2 | T09528 | probable cysteine |
| 86 | 9 | 2.2 | 363 | 2 | S11862 | cysteine proteinase |
| 87 | 9 | 2.2 | 363 | 2 | T06726 | cysteine proteinase |
| 88 | 9 | 2.2 | 364 | 2 | T07707 | cysteine proteinase |
| 89 | 9 | 2.2 | 365 | 2 | T06208 | cysteine proteinase |
| 90 | 9 | 2.2 | 365 | 2 | T06206 | probable cysteine |
| 91 | 9 | 2.2 | 368 | 2 | T01694 | cysteine proteinase |
| 92 | 9 | 2.2 | 371 | 2 | J01111 | cysteine proteinase |
| 93 | 9 | 2.2 | 373 | 2 | J01110 | cysteine proteinase |
| 94 | 9 | 2.2 | 374 | 2 | T06276 | cysteine proteinase |
| 95 | 9 | 2.2 | 374 | 2 | T07541 | leucorhodizole-1 |
| 96 | 9 | 2.2 | 376 | 2 | B96435 | cysteine proteinase |
| 97 | 9 | 2.2 | 377 | 2 | T12042 | cysteine proteinase |
| 98 | 9 | 2.2 | 378 | 2 | S17144 | cysteine proteinase |
| 99 | 9 | 2.2 | 380 | 2 | S55923 | cysteine proteinase |
| 100 | 9 | 2.2 | 436 | 2 | A61377 | probable amino aci |
| 101 | 9 | 2.2 | 436 | 2 | B81034 | acid hydrolase |
| 102 | 9 | 2.2 | 441 | 2 | A45565 | cysteine proteinase |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|---|--------|-----------------------|-----|---|-----|-----|---|--------|---------------------|
| 249 | 6 | 1.5 | 65 | 2 | H85725 | hypothetical prote | 322 | 6 | 1.5 | 146 | 2 | 151473 | cathepsin B-like pr |
| 250 | 6 | 1.5 | 67 | 2 | S10555 | glutathione transp | 323 | 6 | 1.5 | 146 | 2 | A60962 | probable ankyrin |
| 251 | 6 | 1.5 | 67 | 2 | S10555 | hypothetical prote | 324 | 6 | 1.5 | 146 | 2 | G81720 | conserved hypotet |
| 252 | 6 | 1.5 | 72 | 2 | G80951 | hypothetical prote | 325 | 6 | 1.5 | 146 | 2 | A60962 | conserved hypotet |
| 253 | 6 | 1.5 | 72 | 2 | G80951 | hypothetical prote | 326 | 6 | 1.5 | 146 | 2 | G84480 | 1-aminocyclopropan |
| 254 | 6 | 1.5 | 73 | 2 | S27849 | hypothetical prote | 327 | 6 | 1.5 | 148 | 2 | E70366 | conserved hypotet |
| 255 | 6 | 1.5 | 75 | 2 | A31798 | transferrin prote | 328 | 6 | 1.5 | 149 | 2 | S73851 | conserved hypotet |
| 256 | 6 | 1.5 | 76 | 2 | S26474 | hypothetical prote | 329 | 6 | 1.5 | 151 | 2 | A05943 | adhesin p1 hemorag |
| 257 | 6 | 1.5 | 76 | 2 | S58753 | hypothetical prote | 330 | 6 | 1.5 | 151 | 2 | F95316 | probable synth like |
| 258 | 6 | 1.5 | 76 | 2 | S47205 | hypothetical prote | 331 | 6 | 1.5 | 151 | 2 | G95369 | SyB2 transcrip |
| 259 | 6 | 1.5 | 76 | 2 | S47205 | hypothetical prote | 332 | 6 | 1.5 | 152 | 2 | G72642 | hypothetical prote |
| 260 | 6 | 1.5 | 84 | 2 | S44135 | hypothetical prote | 333 | 6 | 1.5 | 153 | 2 | T08564 | hypothetical prote |
| 261 | 6 | 1.5 | 84 | 2 | S44135 | hypothetical prote | 334 | 6 | 1.5 | 153 | 2 | E71226 | hypothetical prote |
| 262 | 6 | 1.5 | 92 | 2 | G28448 | hypothetical prote | 335 | 6 | 1.5 | 154 | 2 | T08500 | hypothetical prote |
| 263 | 6 | 1.5 | 94 | 2 | AH2490 | hypothetical prote | 336 | 6 | 1.5 | 154 | 2 | D08516 | hypothetical prote |
| 264 | 6 | 1.5 | 95 | 2 | F00650 | senescence-associated | 337 | 6 | 1.5 | 155 | 2 | A05606 | hypothetical prote |
| 265 | 6 | 1.5 | 99 | 1 | R6MX12 | ribosomal protein | 338 | 6 | 1.5 | 155 | 2 | G90796 | TEOM protein (impo |
| 266 | 6 | 1.5 | 102 | 2 | G01159 | hypothetical prote | 339 | 6 | 1.5 | 155 | 2 | F31163 | hypothetical prote |
| 267 | 6 | 1.5 | 104 | 2 | A03367 | hypothetical prote | 340 | 6 | 1.5 | 157 | 2 | G01379 | hypothetical prote |
| 268 | 6 | 1.5 | 104 | 2 | G04346 | hypothetical prote | 341 | 6 | 1.5 | 157 | 2 | E96689 | hypothetical prote |
| 269 | 6 | 1.5 | 106 | 2 | S48342 | E7 protein - human | 342 | 6 | 1.5 | 157 | 2 | H81222 | hypothetical prote |
| 270 | 6 | 1.5 | 106 | 2 | H90482 | transcript map w pr | 343 | 6 | 1.5 | 157 | 2 | A12946 | hypothetical prote |
| 271 | 6 | 1.5 | 107 | 2 | A75031 | tau ribosomal prot | 344 | 6 | 1.5 | 157 | 2 | F06722 | hypothetical prote |
| 272 | 6 | 1.5 | 107 | 2 | A83524 | hypothetical prote | 345 | 6 | 1.5 | 157 | 2 | T13296 | hypothetical prote |
| 273 | 6 | 1.5 | 108 | 2 | S45958 | probable membrane | 346 | 6 | 1.5 | 157 | 2 | T48582 | hypothetical prote |
| 274 | 6 | 1.5 | 110 | 2 | G76605 | Qwee spliced varia | 347 | 6 | 1.5 | 157 | 2 | A61342 | hypothetical prote |
| 275 | 6 | 1.5 | 111 | 2 | F71216 | probable ribosomal | 348 | 6 | 1.5 | 158 | 2 | E69456 | hypothetical prote |
| 276 | 6 | 1.5 | 111 | 2 | A62942 | conserved hypotet | 349 | 6 | 1.5 | 158 | 2 | T18107 | hypothetical prote |
| 277 | 6 | 1.5 | 111 | 2 | G98319 | hypothetical prote | 350 | 6 | 1.5 | 158 | 2 | R63555 | hypothetical prote |
| 278 | 6 | 1.5 | 115 | 2 | G98319 | conserved hypotet | 351 | 6 | 1.5 | 157 | 2 | G82115 | hypothetical prote |
| 279 | 6 | 1.5 | 116 | 2 | D49339 | cell growth regula | 352 | 6 | 1.5 | 159 | 2 | F98345 | hypothetical prote |
| 280 | 6 | 1.5 | 116 | 2 | G86129 | probable growth in | 353 | 6 | 1.5 | 159 | 2 | E67428 | hypothetical prote |
| 281 | 6 | 1.5 | 116 | 2 | G01279 | hypothetical prote | 354 | 6 | 1.5 | 159 | 2 | A83622 | hypothetical prote |
| 282 | 6 | 1.5 | 118 | 2 | A01768 | hypothetical prote | 355 | 6 | 1.5 | 161 | 2 | S25305 | hypothetical prote |
| 283 | 6 | 1.5 | 120 | 2 | F71980 | hypothetical prote | 356 | 6 | 1.5 | 161 | 2 | D72682 | hypothetical prote |
| 284 | 6 | 1.5 | 120 | 2 | A64480 | hypothetical prote | 357 | 6 | 1.5 | 162 | 2 | G97548 | hypothetical prote |
| 285 | 6 | 1.5 | 121 | 2 | G90550 | hypothetical prote | 358 | 6 | 1.5 | 162 | 2 | H97261 | hypothetical prote |
| 286 | 6 | 1.5 | 122 | 2 | F84562 | hypothetical prote | 359 | 6 | 1.5 | 162 | 2 | H83215 | hypothetical prote |
| 287 | 6 | 1.5 | 122 | 2 | F49829 | response regulator | 360 | 6 | 1.5 | 163 | 2 | F04186 | hemagglutinin 2 ch |
| 288 | 6 | 1.5 | 124 | 2 | A60288 | hypothetical prote | 361 | 6 | 1.5 | 163 | 2 | S42786 | hypothetical prote |
| 289 | 6 | 1.5 | 124 | 2 | F36609 | hypothetical prote | 362 | 6 | 1.5 | 166 | 2 | E72205 | hypothetical prote |
| 290 | 6 | 1.5 | 127 | 2 | F72112 | hypothetical prote | 363 | 6 | 1.5 | 166 | 2 | A73611 | hypothetical prote |
| 291 | 6 | 1.5 | 129 | 2 | F67943 | ribosomal protein | 364 | 6 | 1.5 | 166 | 2 | F34752 | hypothetical prote |
| 292 | 6 | 1.5 | 129 | 2 | G84590 | hypothetical prote | 365 | 6 | 1.5 | 166 | 2 | F93475 | hypothetical prote |
| 293 | 6 | 1.5 | 129 | 2 | G90529 | hypothetical prote | 366 | 6 | 1.5 | 166 | 2 | E72415 | hypothetical prote |
| 294 | 6 | 1.5 | 130 | 1 | G99078 | hypothetical prote | 367 | 6 | 1.5 | 167 | 2 | H96094 | hypothetical prote |
| 295 | 6 | 1.5 | 132 | 2 | E72643 | hypothetical prote | 368 | 6 | 1.5 | 167 | 2 | A85023 | hypothetical prote |
| 296 | 6 | 1.5 | 133 | 2 | 152082 | glutathione S-tran | 369 | 6 | 1.5 | 168 | 2 | H80866 | hypothetical prote |
| 297 | 6 | 1.5 | 135 | 1 | H5EAH3 | histone H3 - Allen | 370 | 6 | 1.5 | 169 | 2 | T26271 | hypothetical prote |
| 298 | 6 | 1.5 | 135 | 2 | A64828 | histone H3 (family | 371 | 6 | 1.5 | 171 | 2 | G96217 | hypothetical prote |
| 299 | 6 | 1.5 | 136 | 2 | S04099 | histone H3 (family | 372 | 6 | 1.5 | 172 | 2 | H82842 | hypothetical prote |
| 300 | 6 | 1.5 | 136 | 2 | S04099 | histone H3 (family | 373 | 6 | 1.5 | 173 | 2 | A05641 | hypothetical prote |
| 301 | 6 | 1.5 | 136 | 2 | S57626 | histone H3 male | 374 | 6 | 1.5 | 173 | 2 | F97105 | hypothetical prote |
| 302 | 6 | 1.5 | 136 | 2 | S57626 | histone H3 female | 375 | 6 | 1.5 | 174 | 2 | T41632 | hypothetical prote |
| 303 | 6 | 1.5 | 137 | 2 | H59480 | conserved hypotet | 376 | 6 | 1.5 | 175 | 2 | JH0462 | phosphoprotein pho |
| 304 | 6 | 1.5 | 137 | 2 | H84749 | conserved hypotet | 377 | 6 | 1.5 | 175 | 2 | A73352 | hypothetical prote |
| 305 | 6 | 1.5 | 137 | 2 | F97329 | hypothetical prote | 378 | 6 | 1.5 | 175 | 2 | G72683 | hypothetical prote |
| 306 | 6 | 1.5 | 138 | 2 | E90125 | 40S ribosomal prot | 379 | 6 | 1.5 | 175 | 2 | F28850 | hypothetical prote |
| 307 | 6 | 1.5 | 140 | 1 | F28010 | photosystem II pro | 380 | 6 | 1.5 | 175 | 2 | AH1589 | hypothetical prote |
| 308 | 6 | 1.5 | 140 | 1 | G04409 | photosystem II pro | 381 | 6 | 1.5 | 177 | 2 | G71469 | hypothetical prote |
| 309 | 6 | 1.5 | 141 | 2 | T01470 | cystatin S precurs | 382 | 6 | 1.5 | 177 | 2 | A41595 | hypothetical prote |
| 310 | 6 | 1.5 | 142 | 2 | A82517 | ribonuclease P XP2 | 383 | 6 | 1.5 | 178 | 2 | S29483 | hist protein - yem |
| 311 | 6 | 1.5 | 144 | 2 | F92605 | centugal transfer | 384 | 6 | 1.5 | 178 | 2 | S53039 | hypothetical prote |
| 312 | 6 | 1.5 | 144 | 2 | A03572 | hypothetical prote | 385 | 6 | 1.5 | 180 | 2 | G84448 | hypothetical prote |
| 313 | 6 | 1.5 | 144 | 2 | T46379 | hypothetical prote | 386 | 6 | 1.5 | 182 | 2 | KX15 | hypothetical prote |
| 314 | 6 | 1.5 | 145 | 1 | WQ503M | phosphotransferase | 387 | 6 | 1.5 | 182 | 2 | T41204 | hypothetical prote |
| 315 | 6 | 1.5 | 145 | 1 | A41947 | exocytotic synth | 388 | 6 | 1.5 | 182 | 2 | T06972 | hypothetical prote |
| 316 | 6 | 1.5 | 146 | 1 | HDMKH | hemoglobin delta c | 389 | 6 | 1.5 | 183 | 2 | A87344 | hypothetical prote |
| 317 | 6 | 1.5 | 146 | 1 | HDMKTM | hemoglobin delta c | 390 | 6 | 1.5 | 183 | 2 | G65097 | hypothetical prote |
| 318 | 6 | 1.5 | 146 | 1 | HDMKTB | hemoglobin delta c | 391 | 6 | 1.5 | 183 | 2 | D91125 | hypothetical prote |
| 319 | 6 | 1.5 | 146 | 1 | HDMKDU | hemoglobin delta c | 392 | 6 | 1.5 | 183 | 2 | G89370 | hypothetical prote |
| 320 | 6 | 1.5 | 146 | 1 | HDMKSO | hemoglobin delta c | 393 | 6 | 1.5 | 184 | 2 | D87605 | transcripton facto |
| 321 | 6 | 1.5 | 146 | 2 | HDMKSO | superoxide dismuta | 394 | 6 | 1.5 | 184 | 2 | T38216 | hypothetical prote |

| | | | | | | |
|-----|---|-----|-----|---|--------|---------------------|
| 833 | 6 | 1.5 | 363 | 1 | C72590 | probable hexosyltr |
| 834 | 6 | 1.5 | 363 | 1 | S47192 | MHC class I histoc |
| 835 | 6 | 1.5 | 363 | 2 | S40149 | cysteine proteinas |
| 836 | 6 | 1.5 | 363 | 2 | S07113 | class I histocompa |
| 837 | 6 | 1.5 | 363 | 2 | JH0542 | class I histocompa |
| 838 | 6 | 1.5 | 363 | 2 | S03537 | class I histocompa |
| 839 | 6 | 1.5 | 363 | 2 | S21112 | nitrile reductase |
| 840 | 6 | 1.5 | 364 | 2 | T13022 | drought-inducible |
| 841 | 6 | 1.5 | 364 | 2 | E86456 | unknown protein [i |
| 842 | 6 | 1.5 | 364 | 2 | C75629 | CBP family prote |
| 843 | 6 | 1.5 | 365 | 2 | T17048 | cysteine proteinas |
| 844 | 6 | 1.5 | 365 | 2 | S30150 | cysteine proteinas |
| 845 | 6 | 1.5 | 365 | 2 | A85930 | probable glucanase |
| 846 | 6 | 1.5 | 366 | 2 | T17260 | fructose bisphosph |
| 847 | 6 | 1.5 | 367 | 2 | T12382 | cysteine proteinas |
| 848 | 6 | 1.5 | 367 | 2 | S75836 | hypothetical prote |
| 849 | 6 | 1.5 | 368 | 2 | JH0718 | cysteine proteinas |
| 850 | 6 | 1.5 | 369 | 2 | I35306 | hypothetical prote |
| 851 | 6 | 1.5 | 369 | 2 | A83858 | histidinol-phospha |
| 852 | 6 | 1.5 | 370 | 2 | H94455 | protein p15p12.2 |
| 853 | 6 | 1.5 | 371 | 1 | A75043 | probable hexosyltr |
| 854 | 6 | 1.5 | 371 | 2 | T14023 | drought-inducible |
| 855 | 6 | 1.5 | 371 | 2 | T23652 | hypothetical prote |
| 856 | 6 | 1.5 | 372 | 2 | C96419 | p922-3 (imporced) |
| 857 | 6 | 1.5 | 373 | 2 | D71428 | cysteine proteinas |
| 858 | 6 | 1.5 | 373 | 2 | T32900 | hypothetical prote |
| 859 | 6 | 1.5 | 373 | 2 | A69272 | hypothetical prote |
| 860 | 6 | 1.5 | 373 | 2 | E82570 | chitinase mutase |
| 861 | 6 | 1.5 | 374 | 2 | C81412 | alanine dehydrogen |
| 862 | 6 | 1.5 | 374 | 2 | S07815 | mRNA maturase B3 |
| 863 | 6 | 1.5 | 374 | 2 | C94436 | mannitol-1-phospha |
| 864 | 6 | 1.5 | 374 | 2 | T02652 | probable choline k |
| 865 | 6 | 1.5 | 374 | 2 | C96044 | probable reverse t |
| 866 | 6 | 1.5 | 375 | 2 | F71093 | lysine-ERNA 11a |
| 867 | 6 | 1.5 | 376 | 1 | E47905 | conserved hypobact |
| 868 | 6 | 1.5 | 376 | 2 | C72364 | malic oxidoreduct |
| 869 | 6 | 1.5 | 376 | 2 | S56892 | malic oxidoreduct |
| 870 | 6 | 1.5 | 376 | 2 | D84577 | malic oxidoreduct |
| 871 | 6 | 1.5 | 377 | 2 | T47471 | probable rocca prot |
| 872 | 6 | 1.5 | 378 | 2 | T07897 | cysteine proteinas |
| 873 | 6 | 1.5 | 379 | 2 | B4784 | hypothetical prote |
| 874 | 6 | 1.5 | 380 | 1 | TACR | hypothetical prote |
| 875 | 6 | 1.5 | 381 | 1 | C9027 | actindain (p3.4 |
| 876 | 6 | 1.5 | 383 | 1 | C9027 | actindain (p3.4 |
| 877 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 878 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 879 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 880 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 881 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 882 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 883 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 884 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 885 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 886 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 887 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 888 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 889 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 890 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 891 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 892 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 893 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 894 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 895 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 896 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 897 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 898 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 899 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 900 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 901 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 902 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 903 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 904 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |
| 905 | 6 | 1.5 | 383 | 2 | C90641 | cell division prot |

Gene 127, 221-225, 1993
 A>Title: Nucleotide sequence and expression in *Escherichia coli* of cDNA encoding papain
 A'Reference number: J06533; MID:9327235
 A'Accession: J06534
 A:Molecule type: mRNA
 A'Resides: 1367 <REV>
 A'Cross-references: EMBL: M69477, NID:922660, PDB:0AA4594.1; PDB:422661
 A'Experimental source: fruit and leaf
 C'Genetics:
 A:Gene: P-omega
 C:Superfamily: papain
 C'Keywords: cysteine proteinase; hydrolase
 F:1-11/Domain: signal sequence #status predicted <SIG>
 F:12-12/Domain: propeptide #status predicted <SIG>
 F:13-367/Product: carlabin II #status predicted <MAI>
 F:134-195/Protein: carlabin II #status predicted
 F:154-195/Protein: carlabin II #status predicted
 F:157-291/Active site: Cys, His, Asn #status predicted

Query Match 2.5%; Score 10; DB 2; Length 467;
 Best Local Similarity 100.0%; Pred. No. 0.0067;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSAVA 207
 DB 154 CGSCWAFSAVA 164

RESULT 5

S15844

cathepsin S (EC 3.4.22.27) : bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 12 Feb 1994 #sequence_revision 12 Feb 1994 #text_change 18 Jan 1999

C'Accession: S15844; S23690; S16972; S23959

R:Ritter, A.; Collins, A.; Dolency, I.; Orling, T.; Podolsky, M.; Turk, V.

FEBS Lett. 283, 329-331, 1991

A>Title: The complete amino acid sequence of bovine cathepsin S and a partial sequence

A'Reference number: S15844; MID:91257334

A'Accession: S15844

A:Molecule type: protein

A'Resides: 1-217 <RT>

R:Wiederunders, B.; Broemay, D.; Kitzinger, H.; Kalkbrenner, N.; Fling, A.; Papet, I.; Z.

FEBS Lett. 286, 189-192, 1991

A>Title: Primary structure of bovine cathepsin S. Comparison to cathepsins L, H, B and F

A'Reference number: S16972; MID:91323515

A'Accession: S23680

A>Status: not compared with non-phosphatidyl transferase

A:Molecule type: mRNA

A'Resides: 22-217 <MP>

A'Cross-references: EMBL: M69211; NID:9152814; HDB:AAA475.1; PDB:4172415

A>Note: 143-Pro was also found

A'Accession: S16972

A:Molecule type: protein

A'Resides: 1-28; 48-71; 94-104; 107-131; 146-162; 178-217 <MLD>

C:Superfamily: papain

C'Keywords: cysteine proteinase; hydrolase

F:1-110/22-26; 56-76; 90-158-259/Disulfide bonds: #status predicted

F:25; 144/Active site: Cys, His, Asn #status predicted

Query Match 2.5%; Score 10; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 0.045;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSAVA 206
 DB 22 CGSCWAFSAVA 31

PFSMT 6

S68783

cathepsin L (EC 3.4.22.15) precursor - *Paramedium telraurelia*C:Species: *Paramedium telraurelia*

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20 Jan 2000

C'Accession: S68783; S74314

E:Vachek, H.; Kunitz, S.; Linker, D.; Klump, S.; Gans, V.; Durr, G.; Schuster, G.
 Eur. J. Biochem. 249, 199-205, 1996

A>Title: Cathepsin L is an intracellular and extracellular protease in *Paramecium* m14

A'Accession number: S68783; MID:9421819

A'Accession: S68783

A:Molecule type: mRNA

A'Resides: 1-414 <V06>

A'Cross-references: EMBL: X91754; NID:91404086; PDB:1L0087

A'Accession: S74414

A:Molecule type: protein

A'Resides: 120-129; 181-204 <V06>

C'Genetics:

A:Gene: catL

C:Superfamily: papain

C'Keywords: cysteine proteinase; hydrolase

F:1-11/Domain: signal sequence and propeptide #status predicted <SIG>

F:120-213/Protein: cathepsin L #status experimental <MAI>

F:135-255/282/Active site: Cys, His, Asn #status predicted

Query Match 2.5%; Score 10; DB 2; Length 414;

Best Local Similarity 100.0%; Pred. No. 0.052;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSAVA 206
 DB 132 CGSCWAFSAVA 141

RESULT 7

PFSMT 7

papain (EC 3.4.22.2) precursor - papaya

C:Species: *Carica papaya* L. cv. 'Solo'

C:Date: 24 Apr 1984 #sequence_revision 21 Mar 1992 #text_change 18 Jan 1999

C'Accession: A26466; A00974; A37903

R:Cohen, L.M.; Cohen, V.M.; Dilley, L.C.

Gene 18, 213-227, 1985

A>Title: Cloning and sequencing of papain cDNA

A'Reference number: A26466; MID:87163525

A'Accession: A26466

A:Molecule type: mRNA

A'Resides: 1-345 <V06>

A'Cross-references: EMBL: X5203; NID:9152390; PDB:AM62650.1; PDB:467491

F:1-110/22-26; 56-76; 90-158-259/Disulfide bonds: #status predicted

F:25; 144/Active site: Cys, His, Asn #status predicted

A>Title: The complete amino acid sequence of papain. Additions and corrections.

A'Reference number: A00974; MID:71087849

A'Accession: A00974

A:Molecule type: protein

A'Resides: 1-345 <V06>

Native 210, 929-932, 1968

A>Title: Structure of papain.

A'Reference number: A34155; MID:6906973

A'Content: annotation; X-ray crystallography; 2.8 angstroms

R:Vernet, J.; Tessier, D.C.; Richardson, G.; Laliberté, F.; Kouri, H.; Bell, A.W.;

J. Biol. Chem. 265, 16661-16666, 1990

A>Title: Secretion of functional papain precursor from insect cells. Requirement for

A'Reference number: A37903; MID:9075542

A'Accession: A37903

A:Molecule type: protein

A'Resides: 27-49 <V06>

A'Experimental source: recombinant gene expression in a baculovirus/insect cell system

C:Superfamily: papain

C'Keywords: cysteine proteinase; hydrolase

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-134/Domain: propeptide #status predicted <SIG>

F:134-345/Product: papain #status experimental <MAI>

F:155-195/189-229/267-323/Disulfide bonds: #status experimental

F:158-292, 292-309/Active site: Cys, His, Asn #status experimental

Query Match 2.5%; Score 10; DB 1; Length 345;

Best Local Similarity 100.0%; Pred. No. 0.068;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

N: Alternate names: major for excreted protein (MEP), procathepsin L
C: Species: Homo sapiens (man)
C: Cloned: 10, sep-1991 #sequence: Genbank: J03674-136 #date: January 08, 1992
C: Commission: S01002: P04944: synonym: A04944: synonym: P07113: A04944: synonym: P07011: 1979
R: Gal, S.; Gottesman, M.M.
B: Biochem. J. 253, 401-406, 1988
A: Title: Isolation and sequence of a cDNA for human pro cathepsin L
A: Reference number: S01002: MIMD:883960e
A: Accession: S01002
A: Molecule type: mRNA
A: Residues: 1-333 <GAD>
A: Cross-references: GR: X12451, MID: J29714, PDB: CAA29081.1, MID: 929715
J: Joseph, L.J.; Chan, L.C.; Stamenkovich, D.; Sukhtme, V.P.
J: Clin Invest 81, 1621-1629, 1988
A: Title: Complete nucleotide and deduced amino acid sequences of human and murine propeptides
A: Reference number: A92768: MIMD:88213715
A: Accession: B32333
A: Molecule type: mRNA
A: Residues: 1-333 <JOS>
A: Cross-references: GR: M20496, MID: J890235, PDB: AAA66974.1, MID: q10418
J: Joseph, L.J.; Lapid, S.; Sukhtme, V.
Nucleic Acids Res. 15, 3186, 1987
A: Title: The major pro induced protein in NIH3T3 cells is cathepsin L
A: Reference number: S09065, MIMD: 8778433
A: Accession: S09065
A: Molecule type: mRNA
A: Residues: 113-154 <JOS>
A: Cross-references: EMBL: X60566, MID: J27714, PDB: CAA2907.1, MID: q10418
R: Chuhara, S.S.; Popescu, N.C.; Fox, D.; Flisshman, P.; Gottesman, M.M.; Trow, E.B.
J. Biol. Chem. 266, 1039-1045, 1991
A: Title: Cloning, genomic organization, and chromosomal localization of human cathepsin
A: Reference number: A45043: MIMD: 94123212
A: Accession: A45043
A: Molecule type: DNA
A: Residues: 40-46: 82-96: 130-135: 205-210: 259-264: 299-304 cDNA
A: Cross-references: GR: L06426
A: Note: only exon and splice junctions are shown
R: Ritonja, A.; Popovic, F.; Kohnly, M.; Machleidt, W.; Turk, V.
FEBS Lett. 228, 341-345, 1988
A: Title: Amino acid sequences of the human kidney cathepsins B and L
A: Reference number: S00323: MIMD: 9017635
A: Accession: S00323
A: Molecule type: protein
A: Residues: 134-137: 177-179: 238-239: 257-259: 286-287: 333-335 cDNA
R: Machleidt, W.; Ritonja, A.; Popovic, F.; Kohnly, M.; Eising, G.; Turk, V.; Machleidt, W.
In: Cysteine proteinases and their inhibitors, Turk, V., ed., pp. 3-19. Walter de Gruyter, Berlin, 1990
A: Title: Human cathepsins B, H and L: characterization by amino acid sequences and some
A: Reference number: A27911
A: Accession: B27011
A: Molecule type: protein
A: Residues: X, 115-129, M, 131-133, P, 135-141: 292-307, T57, 310-323: M225
R: Mason, P.W.; Walker, T.F.; Northrop, F.D.
Biochem. J. 240, 373-377, 1986
A: Title: The N terminal amino acid sequences of the heavy and light chains of human cathepsin
A: Reference number: A26059: MIMD: 9112952
A: Accession: A26059
A: Molecule type: protein
A: Residues: 114-147, P, 149-152, Y, 292-333: M255
R: Smith, S.M.; Gottesman, M.M.
J. Biol. Chem. 264, 2049-2056, 1989
A: Title: Activity and deletion analysis of recombinant human cathepsin L expressed in bacteria
A: Reference number: A32583: MIMD: 9206733
A: Contents: annotation
C: Genbank:
A: Gene: GDB: CTSL
A: Cross-references: GNR: 119824: OMIM: 116880
A: Map position: 9q22.1-9q22.2
A: Introns: 4273-833, 1573-1574, 2701-3012
C: Complex: heterodimer of disulfide linked chains produced from a single chain precursor
A: Description: catalyzes hydrolysis of peptide bonds in proteins
A: Pathway: intracellular protein degradation
A: Note: Important role in the lysosomal degradation of proteins

[illegible]

T24387
Probable cysteine protease (EC 3.4.22.-) 20366.7 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date of creation: 1997-09-09
C|Accession: T24387
R|Lloyd, C.
Submitted to the EMBL Data Library, March 1997
A|Reference number: 219885
A|Accession: T24387
A|Status: preliminary, translated from cDNA/EMBL/GenBank
A|Molecule type: DNA
A|Positives: 1347 -w11-
A|Cross-references: EMBL:Y52912; PIRN:CA907275.1; CSDDB:CN000219; CNRPT046.7
A|Experimental source: clone T0346
C|Genetics:
A|Gene: CESP:T03E6.7
A|Map position: 5
A|Introns: 86/73; 114/72; 269/3
C|Superfamily: papain
C|Keywords: cysteine protease; hydrolase
F144; 263; 304; 707; 150 Cys, His, Asn status predicted

| | | | |
|----|-----|-----------|-----|
| QY | 197 | CGSCWAFSA | 205 |
| | | | |
| Db | 141 | CGSCWAFSA | 149 |

| | | | |
|----|-----|-----------|-----|
| QY | 197 | CGSCWAFSA | 205 |
| | | | |
| DB | 144 | CGSCWAFSA | 152 |

submitted to the protein sequence database, January 2001.
A: Accession number: Z23015
A: Accession: 145839
A: Status: preliminary
A: Molecule type: DNA
A: Restrictions: 1341 (K16)
A: Cross-references: EMBL:AI149956
A: Experimental source: Cultivar Colombia, PAT 4100-2815
C: Genes/loci:
A: Map position: 3
A: Introns: 149/1
A: NCBI: F0815_200
C: Superfamily: papain
C: Key words: cysteine proteinase, hydrolase
F: 11,284,402/Active site, 678; His, Asp residues predicted

Query Match 2.28; Score 92; 108 Z; Length 341;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
Q7 209 GMAFSAVAA 208
DE 153 GMAFSAVAA 159

[illegible]

RESULT 27
 145839
 Probable cysteine protease (EG 3.4.22) (similarity) - Arabidopsis thaliana
 Nucleotide bound protein FET-20
 Species: Arabidopsis thaliana (mouse-ear cress)
 CDate: 04-Feb-2006 #sequence_classification next_steps 04 Mar 2006
 CAccession: T45839
 P: Ploger, M.; Stahl, C.; Muller-Auer, S.; Schaefer, M.; Wipf, M.; Moos, H.W.; Lemke,

[illegible]

Db 145 CGSCWAFSA 153

RESULT 34

S25267

Cysteine proteinase (EC 3.4.22.1) precursor *Leishmania mexicana*

C:Species: *Leishmania mexicana*

C:Date: 28 May 1993 #sequence_revision 28 May 1993 #text_change 04 Feb 2000

C:Accession: S25267; S16856

R:Mottram, J.C.; Robertson, P.D.; Cowles, G.H.; Barry, J.P.

Mol Microbiol 6, 1995:1992, 1992

AltTitle: A developmentally regulated cysteine proteinase gene of *Leishmania mexicana*.

A:Reference number: S25267; M010:92374843

A:Accession: S25267

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-354 <MOT>

A:Cross-references: EMBL: X62163; NID:94572; UPI:94572

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:153,289,309/Active site: Cys, His, Asn #status predicted

Query Match 2.2% Score 9; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CGSCWAFSA 205

Db 150 CGSCWAFSA 158

RESULT 35

T12039

Cysteine proteinase (EC 3.4.22.1) precursor - kidney bean

C:Species: *Phaseolus vulgaris* (L. Hey. bean)

C:Date: 16-Jul-1999 #sequence_revision 16 Jul 1999 #text_change 20 Jan 2000

C:Accession: T12039

R:Senyuk, V.; Becker, G.; Mertz, K.

Submitted to the EMBL Data Library, October 1997

A:Description: Isolation of cDNA clone encoding cysteine proteinase (CPI) from a cDNA

A:Reference number: Z17385

A:Accession: T12039

A:Status: translated from cDNA/EMBL/GenBank

A:Molecule type: PNA

A:Residues: 1-364 <SEN>

A:Cross-references: EMBL: Z99952

A:Experimental source: cultured Mollicular cell culture clone CP63

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:119/124/Domain: signal sequence #status predicted <SIC>

F:120-124/Domain: propeptide #status predicted <SPRO>

F:125-164/Protein: cysteine proteinase #status predicted <MAT>

F:149,285,305/Active site: Cys, His, Asn #status predicted

Query Match 2.2% Score 9; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 NCOSCWAFSA 204

Db 145 NCOSCWAFSA 153

RESULT 36

T46630

Cysteine proteinase (EC 3.4.22.1) precursor (similarity) - kidney bean

N:Alternate names: cysteine endopeptidase 1

C:Species: *Phaseolus vulgaris* (kidney bean)

C:Date: 16-Feb-2000 #sequence_revision 16 Feb 2000 #text_change 04 Mar 2000

C:Accession: T46630

R:Schlberg, L.E.; Sussex, I.M.

Submitted to the EMBL Data Library, March 1996

A:Description: Sequence of a cDNA encoding a cysteine proteinase from germinating bean seed

A:Reference number: Z23106

A:Accession: T46630

A:Status: preliminary; translated from cDNA/EMBL/GenBank

A:Molecule type: mRNA

A:Residues: 1-64 <SEN>

A:Cross-references: EMBL: U72773; NID:94572; UPI:94572

A:Experimental source: strain Taylor's Horticultural compound

C:Accession: T46630

A:Gene: CP-1

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:119,289,309/Active site: Cys, His, Asn #status predicted

Query Match 2.2% Score 9; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 196 NCOSCWAFSA 204

Db 145 NCOSCWAFSA 153

RESULT 37

T06529

Cysteine proteinase (EC 3.4.22.1) - garden pea

N:Alternate names: NPH

C:Species: *Pisum sativum* (garden pea)

C:Date: 24-Apr-1999 #sequence_revision 24-Apr-1999 #text_change 11-May-2000

C:Accession: T06529

R:Radalsky, J.V.; Brewin, N.J.

Mol. Plant Microbe Interact. 9, 689-695, 1996

AltTitle: Expression of cysteine proteinase genes in pea nodules de novo and ex vivo

A:Reference number: Z15737; M010:97023947

A:Accession: T06529

A:Status: preliminary; translated from cDNA/EMBL/GenBank

A:Molecule type: mRNA

A:Residues: 1-367 <KAR>

A:Cross-references: EMBL: X62163; NID:94572; UPI:94572

A:Experimental source: cv. Wisconsin Perfection, nodules

C:Gene: CPI

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:145,281,301/Active site: Cys, His, Asn #status predicted

Query Match 2.2% Score 9; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 CWASWAFSA 208

Db 145 CWASWAFSA 153

RESULT 38

S47312

Cysteine proteinase (EC 3.4.22.1) precursor - spring vetch

C:Species: *Vicia sativa* (spring vetch, Laro)

C:Date: 06 Jan 1995 #sequence_revision 06 Jan 1995 #text_change 04-Feb-2000

C:Accession: S51817; S47312

R:Becker, G.; Fischer, J.; Nong, V.H.; Munnitz, K.

Plant Mol. Biol. 26, 1207-1212, 1994

AltTitle: PCR cloning and expression analysis of cDNA encoding cysteine proteinases 1

A:Reference number: S51817; M010:95111103

A:Accession: S51817

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-368 <BE2>

A:Cross-references: EMBL: X55739; NID:94572; UPI:94572

AltTitle: the authors translated the cDNA for residue 241 as Gln

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

F:119,289,309/Active site: Cys, His, Asn #status predicted

C:Keywords: cysteine protease; alveolar; hydrolyse; seed
 F1:1-21/Domain: signal sequence status predicted -sig
 F1:22-128/Domain: signal sequence status predicted -sig
 F1:129-344/Protein: 67/100 alpha helices predicted -MAT
 F1:349-458/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:459-192/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:150-192/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:153-286/Active site: Cys, His, Asp #status predicted
 F1:45/Binding site: carboxylate (Asp) (predicted) #status predicted

Query Match 2.2% Score 97 DB 15 Length 458
 Best Local Similarity 100.0% Pred. No. 0.92
 Matches 97 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 197 GSCNAFSA 205
 Db 150 GSCNAFSA 158

RESULT 44

S24602
 Cysteine proteinase tpp (EC 3.4.22.-) - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 20 Feb 1995 #sequence_revision: 20-Feb-1995 #taxi_change: 01 Feb 2000
 C:Accession: S24602
 R:Disbarro, A.G.
 Submitted to the FMBL Data Library, May 1992
 A:Accession: S24602
 A:Reference number: S24602
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-464 <PIS>
 A:Cross-references: EMBL:X66061
 C:Genetics:
 A:Gene: tpp
 A:Start codon: TTG
 C:Superfamily: papain
 C:Keywords: cysteine proteinase; hydrolyse
 F1:1-26/Active site: Cys, His, Asp #status predicted

Query Match 2.2% Score 97 DB 2 Length 464
 Best Local Similarity 100.0% Pred. No. 0.93
 Matches 97 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 197 GSCNAFSA 205
 Db 157 GSCNAFSA 165

RESULT 45

A60667
 Cysteine proteinase cruzain (EC 3.4.22.-) precursor Trypanosoma cruzi
 N:Alternate names: gp57/51
 C:Species: Trypanosoma cruzi
 C:Date: 24-Apr-1994 #sequence_revision: 28 May 1994 #taxi_change: 22-Jul-1994
 C:Accession: A60667; A12497
 R:Murta, A.M.; Persechini, P.M.; de Souza, W.; Guimarães, J.A.; S
 Moll, Biochem Parasitol 43, 27-38, 1990
 Moll, Biochem Parasitol 43, 27-38, 1990
 Moll, Biochem Parasitol 43, 27-38, 1990
 A:Title: Structural and functional identification of gp57/51 antigen in Trypanosoma cruzi
 A:Reference number: A60667; MIM:91148643
 A:Accession: A60667
 A:Molecule type: protein
 A:Residues: 123-125, XX, 131, 133-147, X, 149-152, MGR
 A:Experimental source: strain Y
 R:Bakin, A.E.; Mills, A.A.; Harth, G.; McKerrow, J.H.; Craik, G.S.
 J. Biol. Chem. 267, 7411-7420, 1992
 A:Title: The sequence, organization, and expression of the major cysteine protease (cruzain)
 A:Reference number: A42487; MIM:92218393
 A:Accession: A42487
 A:Molecule type: tRNA
 A:Residues: 1-467 <EAK>
 A:Cross-references: GB:M64342; NID:3162047; FID:AAA20181.1; FID:3162048
 A:Experimental source: strain Tulahuen
 A:Note: sequence extracted from NCBI database (NCBI:09374, NCBI:93747)

C:Superfamily: Papain

C:Keywords: cysteine proteinase; alveolar; hydrolyse
 F1:1-21/Domain: signal sequence status predicted -sig
 F1:22-128/Domain: signal sequence status predicted -sig
 F1:129-344/Protein: 67/100 alpha helices predicted -MAT
 F1:349-458/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:459-192/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:150-192/Protein: carboxyl-terminal propeptide status predicted -MAT
 F1:153-286/Active site: Cys, His, Asp #status predicted
 F1:45/Binding site: carboxylate (Asp) (predicted) #status predicted

Query Match 2.2% Score 97 DB 2 Length 457
 Best Local Similarity 100.0% Pred. No. 0.93
 Matches 97 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 197 GSCNAFSA 205
 Db 144 GSCNAFSA 152

RESULT 46

A45629
 Cysteine proteinase cruzain (EC 3.4.22.-) Trypanosoma cruzi
 C:Species: Trypanosoma cruzi
 C:Date: 22 Apr 1992 #sequence_revision: 19 Nov 1994 #taxi_change: 04-Feb-2000
 C:Accession: A45629
 R:Campbell, C.; Harth, G.; de Souza, W.; Craik, G.S.; McKerrow, J.H.; Buzello
 Moll, Biochem Parasitol 50, 225-234, 1992
 Moll, Biochem Parasitol 50, 225-234, 1992
 A:Title: The major cysteine proteinase (cruzain) from Trypanosoma cruzi is encoded
 A:Reference number: A45629; MIM:92158002
 A:Accession: A45629
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-467 <GMA>
 A:Experimental source: strain Tulahuen 2
 A:Note: sequence extracted from NCBI database (NCBI:M6274, NCBI:M6274)
 C:Superfamily: papain
 C:Keywords: cysteine proteinase; hydrolyse
 F1:1-26/Active site: Cys, His, Asp #status predicted

Query Match 2.2% Score 97 DB 2 Length 473
 Best Local Similarity 100.0% Pred. No. 0.94
 Matches 97 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 197 GSCNAFSA 205
 Db 144 GSCNAFSA 152

RESULT 47

T01206
 Cysteine proteinase m12 (EC 3.4.22.-) m12
 C:Species: Zea mays (maize)
 C:Date: 12 Feb 1999 #sequence_revision: 12 Feb 1999 #taxi_change: 24 Mar 2001
 C:Accession: T01206
 R:Park, Y.; Jia, B.; Yu, L.; Shukla, D.; Bhat, D.; Williams, W.L.
 Submitted to the FMBL Data Library, August 1997
 A:Title: Protein structure of m12 encoded cysteine proteinase from corn (Zea mays L.) callus
 A:Reference number: T01206
 A:Accession: T01206
 A:Molecule type: translated from ORF/EMBL/GenBank
 A:Molecule type: mRNA
 A:Residues: 1-493 <M2>
 A:Cross-references: EMBL:A019145; NID:342593; FID:AAA08242.1; FID:342593
 A:Experimental source: strain Mp7006 callus
 C:Genetics:
 A:Gene: m12
 C:Superfamily: papain
 C:Keywords: cysteine proteinase; hydrolyse
 F1:188-321/Active site: Cys, His, Asp #status predicted
 F1:188-321/Active site: Cys, His, Asp #status predicted

Query Match 2.2% Score 97 DB 2 Length 493
 Best Local Similarity 100.0% Pred. No. 0.94
 Matches 97 Conservative 0 Mismatches 0 Indels 0 Gaps 0

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 9, 2002, 11:46:10, Search time 12 seconds

(without alignments)
1293.878 Million cell updates/sec

Title: US-09-598-062-4

Percent score: 401

Sequence: 1 MDIGNVHEHGVISAPYA.....KKKTCCTIVHVVYVIMQSI 491

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 12 | 3.0 | 346 | 1 | CYST_HUMAN |
| 2 | 11 | 2.7 | 346 | 1 | CYST_HUMAN |
| 3 | 11 | 2.7 | 348 | 1 | PAP3_CARPA |
| 4 | 10 | 2.5 | 217 | 1 | CATS_BOVIN |
| 5 | 10 | 2.5 | 345 | 1 | PAPA_CARPA |
| 6 | 10 | 2.5 | 444 | 1 | TPR-TPRM |
| 7 | 10 | 2.5 | 444 | 1 | CYST_LEP1 |
| 8 | 10 | 2.5 | 471 | 1 | GRV-GRVA |
| 9 | 9 | 2.2 | 217 | 1 | CATL_SHUPP |
| 10 | 9 | 2.2 | 321 | 1 | CYST_HUMAN |
| 11 | 9 | 2.2 | 330 | 1 | CATS_RAT |
| 12 | 9 | 2.2 | 333 | 1 | CATL_HUMAN |
| 13 | 9 | 2.2 | 334 | 1 | CATL_BOVIN |
| 14 | 9 | 2.2 | 334 | 1 | CATL_MOUSE |
| 15 | 9 | 2.2 | 334 | 1 | CATL_PIG |
| 16 | 9 | 2.2 | 334 | 1 | CATL_FAT |
| 17 | 9 | 2.2 | 334 | 1 | CATL_HUMAN |
| 18 | 9 | 2.2 | 354 | 1 | CYST_HUMAN |
| 19 | 9 | 2.2 | 354 | 1 | CYST_HUMAN |
| 20 | 9 | 2.2 | 371 | 1 | CYST_HUMAN |
| 21 | 9 | 2.2 | 458 | 1 | GRV-GRVA |
| 22 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 23 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 24 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 25 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 26 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 27 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 28 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 29 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 30 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 31 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 32 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |
| 33 | 9 | 2.2 | 467 | 1 | CYST_HUMAN |

| | | | | | |
|-----|---|-----|-----|---|------------|
| 34 | 8 | 2.0 | 422 | 1 | CYST_HUMAN |
| 35 | 8 | 2.0 | 423 | 1 | CYST_HUMAN |
| 36 | 8 | 2.0 | 424 | 1 | CYST_HUMAN |
| 37 | 8 | 2.0 | 425 | 1 | CYST_HUMAN |
| 38 | 8 | 2.0 | 426 | 1 | CYST_HUMAN |
| 39 | 8 | 2.0 | 427 | 1 | CYST_HUMAN |
| 40 | 8 | 2.0 | 428 | 1 | CYST_HUMAN |
| 41 | 8 | 2.0 | 429 | 1 | CYST_HUMAN |
| 42 | 8 | 2.0 | 430 | 1 | CYST_HUMAN |
| 43 | 8 | 2.0 | 431 | 1 | CYST_HUMAN |
| 44 | 8 | 2.0 | 432 | 1 | CYST_HUMAN |
| 45 | 8 | 2.0 | 433 | 1 | CYST_HUMAN |
| 46 | 8 | 2.0 | 434 | 1 | CYST_HUMAN |
| 47 | 8 | 2.0 | 435 | 1 | CYST_HUMAN |
| 48 | 8 | 2.0 | 436 | 1 | CYST_HUMAN |
| 49 | 8 | 2.0 | 437 | 1 | CYST_HUMAN |
| 50 | 8 | 2.0 | 438 | 1 | CYST_HUMAN |
| 51 | 8 | 2.0 | 439 | 1 | CYST_HUMAN |
| 52 | 8 | 2.0 | 440 | 1 | CYST_HUMAN |
| 53 | 8 | 2.0 | 441 | 1 | CYST_HUMAN |
| 54 | 8 | 2.0 | 442 | 1 | CYST_HUMAN |
| 55 | 8 | 2.0 | 443 | 1 | CYST_HUMAN |
| 56 | 8 | 2.0 | 444 | 1 | CYST_HUMAN |
| 57 | 8 | 2.0 | 445 | 1 | CYST_HUMAN |
| 58 | 8 | 2.0 | 446 | 1 | CYST_HUMAN |
| 59 | 8 | 2.0 | 447 | 1 | CYST_HUMAN |
| 60 | 8 | 2.0 | 448 | 1 | CYST_HUMAN |
| 61 | 8 | 2.0 | 449 | 1 | CYST_HUMAN |
| 62 | 8 | 2.0 | 450 | 1 | CYST_HUMAN |
| 63 | 8 | 2.0 | 451 | 1 | CYST_HUMAN |
| 64 | 8 | 2.0 | 452 | 1 | CYST_HUMAN |
| 65 | 8 | 2.0 | 453 | 1 | CYST_HUMAN |
| 66 | 8 | 2.0 | 454 | 1 | CYST_HUMAN |
| 67 | 8 | 2.0 | 455 | 1 | CYST_HUMAN |
| 68 | 8 | 2.0 | 456 | 1 | CYST_HUMAN |
| 69 | 8 | 2.0 | 457 | 1 | CYST_HUMAN |
| 70 | 8 | 2.0 | 458 | 1 | CYST_HUMAN |
| 71 | 8 | 2.0 | 459 | 1 | CYST_HUMAN |
| 72 | 8 | 2.0 | 460 | 1 | CYST_HUMAN |
| 73 | 8 | 2.0 | 461 | 1 | CYST_HUMAN |
| 74 | 8 | 2.0 | 462 | 1 | CYST_HUMAN |
| 75 | 8 | 2.0 | 463 | 1 | CYST_HUMAN |
| 76 | 8 | 2.0 | 464 | 1 | CYST_HUMAN |
| 77 | 8 | 2.0 | 465 | 1 | CYST_HUMAN |
| 78 | 8 | 2.0 | 466 | 1 | CYST_HUMAN |
| 79 | 8 | 2.0 | 467 | 1 | CYST_HUMAN |
| 80 | 8 | 2.0 | 468 | 1 | CYST_HUMAN |
| 81 | 8 | 2.0 | 469 | 1 | CYST_HUMAN |
| 82 | 8 | 2.0 | 470 | 1 | CYST_HUMAN |
| 83 | 8 | 2.0 | 471 | 1 | CYST_HUMAN |
| 84 | 8 | 2.0 | 472 | 1 | CYST_HUMAN |
| 85 | 8 | 2.0 | 473 | 1 | CYST_HUMAN |
| 86 | 8 | 2.0 | 474 | 1 | CYST_HUMAN |
| 87 | 8 | 2.0 | 475 | 1 | CYST_HUMAN |
| 88 | 8 | 2.0 | 476 | 1 | CYST_HUMAN |
| 89 | 8 | 2.0 | 477 | 1 | CYST_HUMAN |
| 90 | 8 | 2.0 | 478 | 1 | CYST_HUMAN |
| 91 | 8 | 2.0 | 479 | 1 | CYST_HUMAN |
| 92 | 8 | 2.0 | 480 | 1 | CYST_HUMAN |
| 93 | 8 | 2.0 | 481 | 1 | CYST_HUMAN |
| 94 | 8 | 2.0 | 482 | 1 | CYST_HUMAN |
| 95 | 8 | 2.0 | 483 | 1 | CYST_HUMAN |
| 96 | 8 | 2.0 | 484 | 1 | CYST_HUMAN |
| 97 | 8 | 2.0 | 485 | 1 | CYST_HUMAN |
| 98 | 8 | 2.0 | 486 | 1 | CYST_HUMAN |
| 99 | 8 | 2.0 | 487 | 1 | CYST_HUMAN |
| 100 | 8 | 2.0 | 488 | 1 | CYST_HUMAN |
| 101 | 8 | 2.0 | 489 | 1 | CYST_HUMAN |
| 102 | 8 | 2.0 | 490 | 1 | CYST_HUMAN |
| 103 | 8 | 2.0 | 491 | 1 | CYST_HUMAN |
| 104 | 8 | 2.0 | 492 | 1 | CYST_HUMAN |
| 105 | 8 | 2.0 | 493 | 1 | CYST_HUMAN |
| 106 | 8 | 2.0 | 494 | 1 | CYST_HUMAN |

| | | | | | | | | | | | | | | | |
|-----|---|-----|-----|---|------------|--------|-------------|-----|---|-----|------|---|-------------|--------|--------------|
| 545 | 6 | 1.5 | 684 | 1 | RPOC_MARPO | P06273 | marichantia | 618 | 6 | 1.5 | 999 | 1 | DSN3_HUMAN | P12926 | homo sapien |
| 546 | 6 | 1.5 | 686 | 1 | BDP1_YEAST | P35817 | saccharomyc | 619 | 6 | 1.5 | 1005 | 1 | EAU2_HUMAN | O14491 | homo sapien |
| 547 | 6 | 1.5 | 686 | 1 | MEPD_FIG | P47289 | sus scrofa | 620 | 6 | 1.5 | 1015 | 1 | TTA4_EBEB | G09274 | trichoplax |
| 548 | 6 | 1.5 | 686 | 1 | MEPD_RAT | P24155 | rattus norv | 621 | 6 | 1.5 | 1024 | 1 | STAT_EBEB | P41568 | striptharion |
| 549 | 6 | 1.5 | 688 | 1 | MEPD_HUMAN | P52888 | homo sapien | 622 | 6 | 1.5 | 1029 | 1 | ENK1_YEAST | P12998 | saccharomyc |
| 550 | 6 | 1.5 | 692 | 1 | PHLA_ECOLI | P13323 | escherichia | 623 | 6 | 1.5 | 1039 | 1 | YK04_LERHA | P04342 | trichoplax |
| 551 | 6 | 1.5 | 692 | 1 | PHLA_SALTY | P40734 | salmonella | 624 | 6 | 1.5 | 1042 | 1 | SHWA_MOUSE | P12292 | trichoplax |
| 552 | 6 | 1.5 | 696 | 1 | PLPC_PINEH | P52743 | pinus thurb | 625 | 6 | 1.5 | 1042 | 1 | SV_P_EBEB | O1773 | bovifurcula |
| 553 | 6 | 1.5 | 702 | 1 | LONH_ILINI | O95833 | halobacteri | 626 | 6 | 1.5 | 1058 | 1 | PMAL_DICDI | P46674 | dictyostell |
| 554 | 6 | 1.5 | 710 | 1 | IGAA_SALTY | P56721 | salmonella | 627 | 6 | 1.5 | 1060 | 1 | TFPI_EBEB | G10254 | lyticococcus |
| 555 | 6 | 1.5 | 710 | 1 | IGAA_SALTY | O96810 | salmonella | 628 | 6 | 1.5 | 1079 | 1 | TFPI_EBEB | G10254 | lyticococcus |
| 556 | 6 | 1.5 | 711 | 1 | IGAA_ECO57 | P58720 | escherichia | 629 | 6 | 1.5 | 1082 | 1 | APR2_HUMAN | G14447 | homo sapien |
| 557 | 6 | 1.5 | 711 | 1 | IGAA_ECOJ1 | P45800 | escherichia | 630 | 6 | 1.5 | 1088 | 1 | RRH1_EBEB | P17699 | trichoplax |
| 558 | 6 | 1.5 | 715 | 1 | C17R_MYCTN | P76242 | myctophid | 631 | 6 | 1.5 | 1095 | 1 | A-E1_EBEB | O00209 | homo sapien |
| 559 | 6 | 1.5 | 716 | 1 | GEA1_YEAST | P14742 | saccharomyc | 632 | 6 | 1.5 | 1098 | 1 | GRSA_PV_P92 | P14674 | trichoplax |
| 560 | 6 | 1.5 | 716 | 1 | PAIY_RHOTO | P11544 | rhodospirid | 633 | 6 | 1.5 | 1101 | 1 | GRNG_CELDI | G14626 | trichoplax |
| 561 | 6 | 1.5 | 720 | 1 | YH71_SCHPO | G13326 | schizosacch | 634 | 6 | 1.5 | 1102 | 1 | P116_MOUSE | G09418 | mus muscu |
| 562 | 6 | 1.5 | 723 | 1 | MYSH_MOUSE | P12271 | mus muscu | 635 | 6 | 1.5 | 1102 | 1 | TEAA_EBEB | P56418 | thizobact |
| 563 | 6 | 1.5 | 724 | 1 | W231_MIT22 | P22734 | laetia-phi | 636 | 6 | 1.5 | 1132 | 1 | TRF1_EBEB | P14443 | caenorhabd |
| 564 | 6 | 1.5 | 727 | 1 | KP01_HOVIN | O28161 | bos taurus | 637 | 6 | 1.5 | 1105 | 1 | A4M1_MOUSE | G14624 | mus muscu |
| 565 | 6 | 1.5 | 728 | 1 | KP01_MOUSE | P37350 | mus muscu | 638 | 6 | 1.5 | 1125 | 1 | BE3A_MOUSE | G14624 | mus muscu |
| 566 | 6 | 1.5 | 734 | 1 | PSAB_SYNPA | O78507 | quillardia | 639 | 6 | 1.5 | 1129 | 1 | AK11_PAT | G14624 | mus muscu |
| 567 | 6 | 1.5 | 738 | 1 | PSAB_SYNPA | O94619 | synochococ | 640 | 6 | 1.5 | 1149 | 1 | DPOL_ADPEC | G14624 | mus muscu |
| 568 | 6 | 1.5 | 738 | 1 | TPW4_HUMAN | P08592 | homo sapien | 641 | 6 | 1.5 | 1159 | 1 | P114_HUMAN | P08575 | homo sapien |
| 569 | 6 | 1.5 | 741 | 1 | PIPA_ECOJ1 | P06329 | escherichia | 642 | 6 | 1.5 | 1159 | 1 | RRH2_EBEB | P08575 | homo sapien |
| 570 | 6 | 1.5 | 747 | 1 | PSAB_PROMA | O11835 | homo sapien | 643 | 6 | 1.5 | 1176 | 1 | RPE2_EBEB | P08575 | homo sapien |
| 571 | 6 | 1.5 | 754 | 1 | YAL3_SCHPO | O94340 | schizosacch | 644 | 6 | 1.5 | 1187 | 1 | P1NE_HUMAN | G14624 | mus muscu |
| 572 | 6 | 1.5 | 761 | 1 | AD24_MOUSE | O09903 | schizosacch | 645 | 6 | 1.5 | 1189 | 1 | P1NE_MOUSE | G14624 | mus muscu |
| 573 | 6 | 1.5 | 761 | 1 | TOP1_METPA | O97160 | mus muscu | 646 | 6 | 1.5 | 1220 | 1 | LOXH_HUMAN | G14624 | mus muscu |
| 574 | 6 | 1.5 | 761 | 1 | TOP1_METPA | O56046 | methanococ | 647 | 6 | 1.5 | 1226 | 1 | K4A4_MOUSE | G14624 | mus muscu |
| 575 | 6 | 1.5 | 768 | 1 | YH23_HUMAN | O94137 | homo sapien | 648 | 6 | 1.5 | 1229 | 1 | KRPA_MOUSE | G14624 | mus muscu |
| 576 | 6 | 1.5 | 770 | 1 | AP12_YEAST | P21192 | saccharomyc | 649 | 6 | 1.5 | 1239 | 1 | ORNA_EBEB | G14624 | mus muscu |
| 577 | 6 | 1.5 | 783 | 1 | K6PF_ASPIC | P78985 | aspergillus | 650 | 6 | 1.5 | 1279 | 1 | YTRN_EBEB | G14624 | mus muscu |
| 578 | 6 | 1.5 | 787 | 1 | PFL_LACTA | O32797 | lactococcus | 651 | 6 | 1.5 | 1280 | 1 | MYSH_HUMAN | G14624 | mus muscu |
| 579 | 6 | 1.5 | 787 | 1 | PFL_LACTA | O32799 | lactococcus | 652 | 6 | 1.5 | 1279 | 1 | VALN_CHICK | G14624 | mus muscu |
| 580 | 6 | 1.5 | 788 | 1 | SYPR_THPMA | O94259 | thrombocyte | 653 | 6 | 1.5 | 1311 | 1 | ALPX_DROME | G14624 | mus muscu |
| 581 | 6 | 1.5 | 795 | 1 | CHUB_HUMAN | O35511 | homo sapien | 654 | 6 | 1.5 | 1315 | 1 | SEB4_YEAST | G14624 | mus muscu |
| 582 | 6 | 1.5 | 795 | 1 | RPO4_VACCV | P07241 | vaccinia vi | 655 | 6 | 1.5 | 1328 | 1 | HNS2_SCHPO | G14624 | mus muscu |
| 583 | 6 | 1.5 | 805 | 1 | ACVS_STECL | P33697 | variola vir | 656 | 6 | 1.5 | 1330 | 1 | WFS_MOUSE | G14624 | mus muscu |
| 584 | 6 | 1.5 | 805 | 1 | ACVS_STECL | O01757 | streptomyce | 657 | 6 | 1.5 | 1352 | 1 | Y944_HUMAN | G14624 | mus muscu |
| 585 | 6 | 1.5 | 808 | 1 | GC81_SCHPO | O14255 | schizosacch | 658 | 6 | 1.5 | 1367 | 1 | DP03_THPMA | G14624 | mus muscu |
| 586 | 6 | 1.5 | 819 | 1 | A2A1_HUMAN | O95782 | h adapter r | 659 | 6 | 1.5 | 1376 | 1 | RH02_LANAU | P52754 | arabidopsi |
| 587 | 6 | 1.5 | 828 | 1 | HYBH_TREBY | O54316 | triforcea b | 660 | 6 | 1.5 | 1384 | 1 | RH02_SINAI | G14624 | mus muscu |
| 588 | 6 | 1.5 | 832 | 1 | Y256_YEAST | P35611 | saccharomyc | 661 | 6 | 1.5 | 1401 | 1 | LYS2_CANAL | G14624 | mus muscu |
| 589 | 6 | 1.5 | 843 | 1 | AP01_MOUSE | O54312 | mus muscu | 662 | 6 | 1.5 | 1399 | 1 | RPE2_EBEB | G14624 | mus muscu |
| 590 | 6 | 1.5 | 845 | 1 | AMPN_LACTC | P37897 | lactococcus | 663 | 6 | 1.5 | 1409 | 1 | RPE2_PSEB0 | G14624 | mus muscu |
| 591 | 6 | 1.5 | 848 | 1 | PP0P_MYSEC | P47326 | mycoplasma | 664 | 6 | 1.5 | 1411 | 1 | Y197_PSEB0 | G14624 | mus muscu |
| 592 | 6 | 1.5 | 865 | 1 | PM11_HUMAN | O43460 | homo sapien | 665 | 6 | 1.5 | 1413 | 1 | SHUP_PV_P90 | G14624 | mus muscu |
| 593 | 6 | 1.5 | 867 | 1 | EP2_BLAHO | O17152 | blastocysti | 666 | 6 | 1.5 | 1453 | 1 | NRKR_MOUSE | G14624 | mus muscu |
| 594 | 6 | 1.5 | 872 | 1 | DP3A_MYCPN | P75404 | mycoplasma | 667 | 6 | 1.5 | 1501 | 1 | SN02_YEAST | G14624 | mus muscu |
| 595 | 6 | 1.5 | 894 | 1 | NIA_BABBA | P43100 | brucella b | 668 | 6 | 1.5 | 1668 | 1 | DPOL_THPMA | G14624 | mus muscu |
| 596 | 6 | 1.5 | 905 | 1 | NIA_FUSOX | P33863 | insulinum o | 669 | 6 | 1.5 | 1680 | 1 | KETA_HUMAN | G14624 | mus muscu |
| 597 | 6 | 1.5 | 917 | 1 | MSH2_DROME | P43248 | trichoplax | 670 | 6 | 1.5 | 1709 | 1 | CHH1_HUMAN | G14624 | mus muscu |
| 598 | 6 | 1.5 | 918 | 1 | Y438_SCHPO | O14164 | schizosacch | 671 | 6 | 1.5 | 1711 | 1 | CHH1_MOUSE | G14624 | mus muscu |
| 599 | 6 | 1.5 | 919 | 1 | ATC1_HUMAN | P58194 | homo sapien | 672 | 6 | 1.5 | 1722 | 1 | YH04_CAEEL | G14624 | mus muscu |
| 600 | 6 | 1.5 | 919 | 1 | ATC1_PAT | O64566 | rattus norv | 673 | 6 | 1.5 | 1729 | 1 | CHH2_HUMAN | G14624 | mus muscu |
| 601 | 6 | 1.5 | 932 | 1 | HIRE_SCHPO | P87314 | schizosacch | 674 | 6 | 1.5 | 1728 | 1 | MYSA_RAT | G14624 | mus muscu |
| 602 | 6 | 1.5 | 937 | 1 | A2A2_PAT | P14934 | rattus norv | 675 | 6 | 1.5 | 1829 | 1 | MYSA_HUMAN | G14624 | mus muscu |
| 603 | 6 | 1.5 | 938 | 1 | A2A2_MOUSE | P17427 | mus muscu | 676 | 6 | 1.5 | 1846 | 1 | MYSH_RAT | G14624 | mus muscu |
| 604 | 6 | 1.5 | 939 | 1 | A2A2_HUMAN | O94973 | h adapter r | 677 | 6 | 1.5 | 1853 | 1 | MYGA_MOUSE | G14624 | mus muscu |
| 605 | 6 | 1.5 | 947 | 1 | PM2A_APATH | P19456 | arabidopsi | 678 | 6 | 1.5 | 1875 | 1 | MYGA_HUMAN | G14624 | mus muscu |
| 606 | 6 | 1.5 | 949 | 1 | PM6A_APATH | O95876 | arabidopsi | 679 | 6 | 1.5 | 2073 | 1 | FASL_SCHPO | G14624 | mus muscu |
| 607 | 6 | 1.5 | 953 | 1 | ATC1_BOVIN | P57709 | bos taurus | 680 | 6 | 1.5 | 2109 | 1 | KR01_VSVST | G14624 | mus muscu |
| 608 | 6 | 1.5 | 954 | 1 | PM6A_APATH | O15556 | arabidopsi | 681 | 6 | 1.5 | 2116 | 1 | KR01_VSVST | G14624 | mus muscu |
| 609 | 6 | 1.5 | 963 | 1 | ATC4_HUMAN | O75185 | homo sapien | 682 | 6 | 1.5 | 2242 | 1 | AP01_PAT | G14624 | mus muscu |
| 610 | 6 | 1.5 | 970 | 1 | TPA1_MOUSE | O04222 | comptosia | 683 | 6 | 1.5 | 2242 | 1 | AP01_PAT | G14624 | mus muscu |
| 611 | 6 | 1.5 | 971 | 1 | MCM1_YEAST | P24279 | saccharomyc | 684 | 6 | 1.5 | 2345 | 1 | AP01_MOUSE | G14624 | mus muscu |
| 612 | 6 | 1.5 | 974 | 1 | AMP_HUMAN | P19021 | homo sapien | 685 | 6 | 1.5 | 2345 | 1 | HY02_EBEB | G14624 | mus muscu |
| 613 | 6 | 1.5 | 977 | 1 | A2A1_MOUSE | P17426 | mus muscu | 686 | 6 | 1.5 | 4080 | 1 | ZPH2_DROME | G14624 | mus muscu |
| 614 | 6 | 1.5 | 979 | 1 | AML_MOUSE | P97467 | mus muscu | 687 | 6 | 1.5 | 4080 | 1 | P046_ZYWAR | G14624 | mus muscu |
| 615 | 6 | 1.5 | 980 | 1 | PEX6_HUMAN | O13608 | homo sapien | 688 | 6 | 1.5 | 4083 | 1 | P046_ZYWAR | G14624 | mus muscu |
| 616 | 6 | 1.5 | 989 | 1 | TOP_MOUSE | P22817 | trichoplax | 689 | 6 | 1.5 | 4092 | 1 | CTA2_EBEB | G14624 | mus muscu |
| 617 | 6 | 1.5 | 989 | 1 | SPFA_PLAFC | P13823 | plasmodium | 690 | 6 | 1.5 | 4106 | 1 | LYA2_MOUSE | G14624 | mus muscu |

| | | | | | | | | | | | | | | | |
|-----|---|-----|----|---|------------|--------|---------------|-----|---|-----|-----|---|------------|--------|-------------|
| 837 | 5 | 1.2 | 91 | 1 | S112_BV1N | P79105 | bos taurus | 910 | 5 | 1.2 | 103 | 1 | EMBL_H1PK | P79417 | homo sapien |
| 838 | 5 | 1.2 | 91 | 1 | VNS1_AIN-V | P14113 | galea virus | 911 | 5 | 1.2 | 102 | 1 | EMBL_DICD1 | P16648 | dicotyledon |
| 839 | 5 | 1.2 | 91 | 1 | Y606_MCTJA | G58023 | machorococ | 912 | 5 | 1.2 | 103 | 1 | MGF_JWV1N | P29732 | bos taurus |
| 840 | 5 | 1.2 | 91 | 1 | Y614_HABIN | P44186 | haemophilus | 913 | 5 | 1.2 | 103 | 1 | MGF_JWV1N | P04494 | rattus norv |
| 841 | 5 | 1.2 | 92 | 1 | MOT1_PORSE | G46617 | equus caball | 914 | 5 | 1.2 | 103 | 1 | MGF_JWV1N | G47479 | arabidopsi |
| 842 | 5 | 1.2 | 92 | 1 | PTC2_SALTY | G91683 | salmonella | 915 | 5 | 1.2 | 103 | 1 | MGF_JWV1N | P29424 | arabidopsi |
| 843 | 5 | 1.2 | 92 | 1 | PTC1_PORSE | G91683 | salmonella | 916 | 5 | 1.2 | 103 | 1 | MGF_JWV1N | P29424 | arabidopsi |
| 844 | 5 | 1.2 | 92 | 1 | Y606_MCTJA | P16161 | yeast | 917 | 5 | 1.2 | 104 | 1 | CH12_BV1A | P39663 | bradyrhizob |
| 845 | 5 | 1.2 | 93 | 1 | H011_HRP2 | P51773 | bacteriophage | 918 | 5 | 1.2 | 104 | 1 | CH13_BV1A | P39663 | bradyrhizob |
| 846 | 5 | 1.2 | 93 | 1 | PAH6_CANVA | P31299 | canis famli | 919 | 5 | 1.2 | 104 | 1 | CYC_SOURCE | P39663 | bradyrhizob |
| 847 | 5 | 1.2 | 94 | 1 | SCOP_PANIM | P56972 | pandanus im | 920 | 5 | 1.2 | 104 | 1 | RECA_LAGRE | P39663 | bradyrhizob |
| 848 | 5 | 1.2 | 95 | 1 | CH10_PYMO | P19229 | yeast | 921 | 5 | 1.2 | 104 | 1 | RECA_LAGRE | P39663 | bradyrhizob |
| 849 | 5 | 1.2 | 95 | 1 | IMQ2_P1EPM | G91480 | rice | 922 | 5 | 1.2 | 104 | 1 | RECA_LAGRE | P39663 | bradyrhizob |
| 850 | 5 | 1.2 | 95 | 1 | PAH6_HUMAN | P01246 | homo sapien | 923 | 5 | 1.2 | 104 | 1 | RECA_LAGRE | P39663 | bradyrhizob |
| 851 | 5 | 1.2 | 95 | 1 | PEPM_FCOL1 | P31231 | escherichia | 924 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 852 | 5 | 1.2 | 95 | 1 | VE10_VAVOC | P21050 | vacuola | 925 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 853 | 5 | 1.2 | 95 | 1 | VE10_VAVOC | P21050 | vacuola | 926 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 854 | 5 | 1.2 | 95 | 1 | VE10_VAVOC | P21050 | vacuola | 927 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 855 | 5 | 1.2 | 96 | 1 | C553_HED17 | G91480 | rice | 928 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 856 | 5 | 1.2 | 96 | 1 | C553_HED17 | G91480 | rice | 929 | 5 | 1.2 | 104 | 1 | TH10_H0178 | P39663 | bradyrhizob |
| 857 | 5 | 1.2 | 96 | 1 | CH10_NL1G0 | P77913 | neisseria | 930 | 5 | 1.2 | 105 | 1 | W0N1_S0110 | P39663 | bradyrhizob |
| 858 | 5 | 1.2 | 96 | 1 | HYPC_NHOCA | P40428 | rhodobacter | 931 | 5 | 1.2 | 105 | 1 | Y450_A0040 | P39663 | bradyrhizob |
| 859 | 5 | 1.2 | 96 | 1 | PAH6_CANVA | P31299 | canis famli | 932 | 5 | 1.2 | 105 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 860 | 5 | 1.2 | 96 | 1 | R37A_PLAFA | O94184 | plasmid | 933 | 5 | 1.2 | 105 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 861 | 5 | 1.2 | 96 | 1 | BL11_HSV11 | P04289 | herpes simp | 934 | 5 | 1.2 | 105 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 862 | 5 | 1.2 | 96 | 1 | BL11_HSV11 | P04289 | herpes simp | 935 | 5 | 1.2 | 105 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 863 | 5 | 1.2 | 96 | 1 | YEBE_FCOL1 | P31231 | escherichia | 936 | 5 | 1.2 | 106 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 864 | 5 | 1.2 | 96 | 1 | YEBE_FCOL1 | P31231 | escherichia | 937 | 5 | 1.2 | 106 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 865 | 5 | 1.2 | 97 | 1 | YEBE_FCOL1 | P31231 | escherichia | 938 | 5 | 1.2 | 106 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 866 | 5 | 1.2 | 97 | 1 | YEBE_FCOL1 | P31231 | escherichia | 939 | 5 | 1.2 | 106 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 867 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 940 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 868 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 941 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 869 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 942 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 870 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 943 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 871 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 944 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 872 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 945 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 873 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 946 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 874 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 947 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 875 | 5 | 1.2 | 98 | 1 | NGM_CROLA | P34193 | crossostoma | 948 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 876 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 949 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 877 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 950 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 878 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 951 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 879 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 952 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 880 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 953 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 881 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 954 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 882 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 955 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 883 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 956 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 884 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 957 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 885 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 958 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 886 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 959 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 887 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 960 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 888 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 961 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 889 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 962 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 890 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 963 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 891 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 964 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 892 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 965 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 893 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 966 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 894 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 967 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 895 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 968 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 896 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 969 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 897 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 970 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 898 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 971 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 899 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 972 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 900 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 973 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 901 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 974 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 902 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 975 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 903 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 976 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 904 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 977 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 905 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 978 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 906 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 979 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 907 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 980 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 908 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 981 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |
| 909 | 5 | 1.2 | 99 | 1 | NGM_CROLA | P34193 | crossostoma | 982 | 5 | 1.2 | 107 | 1 | Y819_R1018 | P39663 | bradyrhizob |

OY 197 CQSCWAFSAVA 207
 DB 154 CQSCWAFSAVA 164

RESULT 3
 PAP3_CARPA STANDARD: PRI: 348 AA.

DT 01-MAR-1989 (rel. 10, created)
 DT 01-OCT-1996 (rel. 34, last sequence update)
 DT 01-MAR-2002 (rel. 41, last annotation update)
 DE Caricain precursor (for 3.4.22.40) (Papaya fruitainase source) (Papaya
 protease III) (PPIII) (Papaya peptidase A).
 OS Carica papaya (Papaya).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Caricaceae; Carica.
 OX NCBI_taxid:36449;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-Leaf;
 RA Baker K.C., Reveal J.D.F., Cummings N.J., Collins M.E., Goodenough F.W.;
 RA Submitted (MAY 1995) to the EMBL, GenBank, DDBJ databases.
 RN 12
 RP SEQUENCE OF 133-348.
 RX MEDLINE:89105320; PubMed:3063283;
 RA Dubois T., Kleinschmidt T., Schnek A.G., Looze Y., Braunitzer G.;
 RT "The thiol proteases from the latex of Carica papaya L. II. The
 RT primary structure of proteinase omega."
 RL Biol. Chem. Hoppe-Seyler 269:741-754(1988).
 RN 13
 RP SEQUENCE OF 237-348 FROM N.A.
 RC TISSUE-Leaf;
 RA Collins M.E., Reveal J.D.F., Sumner I.G., Pickersgill R.W.;
 RA Goodenough F.W.;
 RA Submitted (FEB-1990) to the EMBL/GenBank/DDBJ databases.
 RN 14
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RA Pickersgill R.W., Rickallah F., Bartle G.W., Goodenough F.W.;
 RT "Determination of the structure of papaya proteinase omega."
 RL Acta Crystallogr. B 47:756-771(1991).
 RN 15
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF MUTANT G10-290.
 RX MEDLINE:96354872; PubMed:8759310;
 RA Katerelos N.A., Taylor M.A.J., Scott M., Goodenough F.W.;
 RA Pickersgill R.W.;
 RT "Crystal structure of a caricain D158E mutant in complex with E-64."
 RL FEBS Lett. 392:335-39(1996).
 RN 16
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF PRO-CARICAIN.
 RX MEDLINE:97094976; PubMed:8919744;
 RA Groves M.P., Taylor M.A., Scott M., Cummings N.J., Pickersgill R.W.;
 RA Jenkins J.A.;
 RT "The prosequence of procaricain forms an alpha-helical domain that
 RT prevents access to the substrate-binding cleft."
 RL Structure 4:1193-1203(1996).
 RN 17
 RP 3D-STRUCTURE MODELING OF 133-348.
 RA Paratcharya S., Pal A., Bera A., Chakrabarty S., Banerjee A.;
 RT "Theoretical model structure of a caricain D158E mutant in complex
 RT with E-64C."
 RI Submitted (APR-2000) to the PDB data bank
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 CC for peptide bonds, similar to those of papain and chymopapain.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAINE FAMILY OF THIOL PROTEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is not
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement with the EMBL Bioinformatics
 CC or send an email to license@ebi.ac.uk.

DR EMBL: X60600; CAA66662.1;
 DR EMBL: X61899; CAA61801.1;
 DR PIR: S01135; S01145;
 DR PIR: S11748; S11748;
 DR PDB: 1P0J; 31-JAN-94.
 DR PDB: 1MEC; 11-JAN-97.
 DR PDB: 1PCL; 01-APR-97.
 DR PDB: 1P0X; 26-JUN-00.
 DR PDB: 1C7X; 03-MAY-00.
 DR MEROPS: C01.0043;
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR001693; Thiolprotease_act_site.
 DR Pfam: PF00112; Peptidase_C1_1.
 DR Pfam: PF07965; PAPAINE.
 DR PROSITE: PS00149; THIOL_PROTEASE_CYS_1;
 DR PROSITE: PS00649; THIOL_PROTEASE_HIS_1;
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN_1;
 KW Hydrolyase; Thiol protease; Zymogen; Signal; 3D-structure;
 FT SIGNAL: 1
 FT CHAIN: 133-348
 FT ACT_SITE: 157
 FT ACT_SITE: 291
 FT ACT_SITE: 311
 FT DISULFID: 154
 FT DISULFID: 188
 FT DISULFID: 227
 FT STRAND: 286
 FT STRAND: 147
 FT TURN: 149
 FT STRAND: 150
 FT STRAND: 155
 FT HELIX: 157
 FT STRAND: 160
 FT HELIX: 182
 FT TURN: 194
 FT STRAND: 196
 FT HELIX: 210
 FT STRAND: 212
 FT STRAND: 214
 FT TURN: 215
 FT TURN: 215
 FT HELIX: 229
 FT STRAND: 247
 FT STRAND: 241
 FT HELIX: 250
 FT STRAND: 262
 FT HELIX: 271
 FT STRAND: 280
 FT STRAND: 291
 FT TURN: 302
 FT STRAND: 304
 FT STRAND: 313
 FT TURN: 418
 FT STRAND: 419
 FT TURN: 420
 FT STRAND: 421
 FT STRAND: 426
 FT HELIX: 435
 FT TURN: 439
 FT STRAND: 444
 FT STRAND: 446
 FT STRAND: 446

Query Match 2.78; Score 11; DR 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 0.0042;
 Matches 11; Conservative 0; Mismatches 0; Gaps 0

OY 197 CQSCWAFSAVA 207
 DB 154 CQSCWAFSAVA 164

| | | |
|----------|---|---|
| R | | cystatin proteinases of rice seeds (oryzatin) " |
| RL | J Biol Chem 267:16897-16902(1992) | |
| CC | -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN SEEDS. | |
| CC | -1- INDUCTION: BY GIBBERELIC ACID (GA). | |
| CC | -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE PAPAIN FAMILY OF THIOL PROTEASES. | |
| CC | | |
| CC | This Swiss-prot entry is deprecated. It is replaced by a collaboration between the Swiss Institute of Bioinformatics and the EMBL consortium at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is to be kept modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. For more information, please send an email to license@sib-swiss.ch). | |
| CC | | |
| DR | EMBL: D90407 : BAA14403.1 ; - | |
| DR | PAP: J00389 ; KHRZOB. | |
| DR | HSSP: P00785 ; ZACTI. | |
| DR | MROPS: C01.029 ; - | |
| DR | Interpro: IPR000118 ; Granulin. | |
| DR | Interpro: IPR000668 ; Peptidase_C1. | |
| DR | Interpro: IPR000169 ; ThiolProt_act_site | |
| DE | Plant_F093395 ; Granulin_1 | |
| DE | Pfam: PF00112 ; peptidase_c1.1. | |
| DR | PRINTS: PR00705 ; PAPAIN. | |
| DR | SMART: SMO0277 ; GRAN_1. | |
| DR | PROSITE: PS00139 ; THIOL_PROTEASE_CYS_1. | |
| DR | PROSITE: PS00639 ; THIOL_PROTEASE_HIS_1. | |
| DP | PROSITE: PS00640 ; THIOL_PROTEASE ASN_1 | |
| KW | Hydrolase, Thiol, Protease, Catalytic, papain-like, catalytic, SIGNAL. | |
| FT | SIGNAL 1 21 POTENTIAL. | |
| FT | PROTER 22 179 ACTIVATION INTERPE (PROBETIAN). | |
| FT | CHAIN 140 471 ORYZAON BETA CHAIN. | |
| FT | ACT_SITE 164 361 BY SIMILARITY. | |
| FT | ACT_SITE 301 301 BY SIMILARITY. | |
| FT | ACT_SITE 321 321 BY SIMILARITY. | |
| FT | DISULFID 161 204 BY SIMILARITY. | |
| FT | DISULFID 195 232 BY SIMILARITY. | |
| FT | DISULFID 295 346 BY SIMILARITY. | |
| FT | CARBOHYD 340 340 N-LINKED (GLCNAC...) (POTENTIAL). | |
| FT | CARBOHYD 388 388 N-LINKED (GLCNAC...) (POTENTIAL). | |
| FT | SPQOBNYD 471 AA: 50505 MW: 09F39D3399SCERHD GRG64; | |
| SO | SOURCE Match 2.5% Score 109.89 Id Length 471;
Best Local Similarity 100.0%; Prod No. 0.642;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0 | |
| QY | 197 CCSCWAFSAV 206

Db 161 CCSCWAFSAV 170 | |
| RESULT 9 | | |
| CD | CATL_SHEEP STANDARD: PIR: 217 AA. | |
| AD | CATL_SHEEP Q10931: | |
| DT | 61-oct-1996 [rel. 34, created] | |
| DT | 01-oct-1996 [rel. 34, last sequence update] | |
| DT | 16-oct-2001 [rel. 40, first time in update] | |
| DE | Cathepsin L (EC 3.4.22.15). | |
| GN | CITSL. | |
| OS | Ovis aries (Sheep). | |
| OC | Eukaryota; Metazoa; Chordata; Gracilaria; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Cetartiodactyla; Artiodactyla; Bovidae; Caprinae; | |
| OC | Ruminantia; Capriinae; Ovis. | |
| OX | NCBI_TaxId=5940; | |
| RN | | |
| PP | SEQUENCE: | |
| RC | TISSUE=Liver; | |
| RX | MEDLINE=96321798; PubMed 8794201; | |
| RA | Ritovska A., Goelzer T.H.T., Pike R.N., Dousson C.J. | |
| RT | "The amino acid sequences, structure comparisons and inhibition kinetics of Sheep cathepsin L and sheep stromal B." | |

| | | |
|-----|-------------------------|--|
| 24 | Comp. Biochem. Physiol. | 118B:13-19(1996). |
| 25 | 1 | Localized, Membrane Protein, the Cytosolic Domain of a Protein in |
| 26 | | Lysosomes. |
| 27 | 1 | CATALYTIC ACTIVITY: Specifically close to that of papain. As |
| 28 | | compared to cathepsin B, cathepsin L exhibits higher activity |
| 29 | | towards protein substrates, but has little activity on Z Arg-Arg |
| 30 | | NHMe, and no peptidyl-dipeptidase activity. |
| 31 | 1 | Specific cleavage of a heavy and a light chain linker by disulfide |
| 32 | | bonds. |
| 33 | 1 | SUBCELLULAR LOCATION: Lysosomal. |
| 34 | 1 | SIMILARITY: BELONGS TO THE LYSOGENE FAMILY. CL. AND FY. AS THE |
| 35 | | PAPAIN FAMILY OF THIOL PROTEASES. |
| 36 | 1 | 100% IDENTITY. |
| 37 | DR | MEKOP1: G01.0327. |
| 38 | DR | InterPro: IPR000668: peptidase_CL. |
| 39 | DR | InterPro: IPR00169: ThiolProteol_act_site. |
| 40 | DR | Plant: PF00112: peptidase_CL_1. |
| 41 | DR | PRINTS: PR00705: PAPAOL. |
| 42 | DR | PROSITE: PS00159: THIOL_PROTEASE_1/MS. |
| 43 | DR | PROSITE: PS00619: THIOL_PROTEASE_HIS_1. |
| 44 | DR | PROSITE: PS00640: THIOL_PROTEASE_ASG_1. |
| 45 | DR | Biological: Thiol protease: glycoprotein: lysosome. |
| 46 | DR | CHAIN: 1 |
| 47 | FT | CHAIN: 127 |
| 48 | FT | ACT_SITE: 25 |
| 49 | FT | ACT_SITE: 163 |
| 50 | FT | ACT_SITE: 184 |
| 51 | FT | DISULFID: 22 |
| 52 | FT | DISULFID: 166 |
| 53 | FT | DISULFID: 206 |
| 54 | FT | DISULFID: 217 |
| 55 | FT | DISULFID: 217 |
| 56 | FT | DISULFID: 217 |
| 57 | FT | DISULFID: 217 |
| 58 | FT | DISULFID: 217 |
| 59 | FT | DISULFID: 217 |
| 60 | FT | DISULFID: 217 |
| 61 | FT | DISULFID: 217 |
| 62 | FT | DISULFID: 217 |
| 63 | FT | DISULFID: 217 |
| 64 | FT | DISULFID: 217 |
| 65 | FT | DISULFID: 217 |
| 66 | FT | DISULFID: 217 |
| 67 | FT | DISULFID: 217 |
| 68 | FT | DISULFID: 217 |
| 69 | FT | DISULFID: 217 |
| 70 | FT | DISULFID: 217 |
| 71 | FT | DISULFID: 217 |
| 72 | FT | DISULFID: 217 |
| 73 | FT | DISULFID: 217 |
| 74 | FT | DISULFID: 217 |
| 75 | FT | DISULFID: 217 |
| 76 | FT | DISULFID: 217 |
| 77 | FT | DISULFID: 217 |
| 78 | FT | DISULFID: 217 |
| 79 | FT | DISULFID: 217 |
| 80 | FT | DISULFID: 217 |
| 81 | FT | DISULFID: 217 |
| 82 | FT | DISULFID: 217 |
| 83 | FT | DISULFID: 217 |
| 84 | FT | DISULFID: 217 |
| 85 | FT | DISULFID: 217 |
| 86 | FT | DISULFID: 217 |
| 87 | FT | DISULFID: 217 |
| 88 | FT | DISULFID: 217 |
| 89 | FT | DISULFID: 217 |
| 90 | FT | DISULFID: 217 |
| 91 | FT | DISULFID: 217 |
| 92 | FT | DISULFID: 217 |
| 93 | FT | DISULFID: 217 |
| 94 | FT | DISULFID: 217 |
| 95 | FT | DISULFID: 217 |
| 96 | FT | DISULFID: 217 |
| 97 | FT | DISULFID: 217 |
| 98 | FT | DISULFID: 217 |
| 99 | FT | DISULFID: 217 |
| 100 | FT | DISULFID: 217 |

[illegible]

FT CHAIN 291 334 CATHEPSIN L, LIGHT CHAIN.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT ACT_SITE 276 276 BY SIMILARITY.
 FT ACT_SITE 300 300 BY SIMILARITY.
 FT DISULFID 145 178 BY SIMILARITY.
 FT DISULFID 169 211 BY SIMILARITY.
 FT DISULFID 269 223 INTERCHAIN (BY SIMILARITY).
 FT CAPSHEID 221 221 N-LINKED (GLYCAN...)
 FT COMPLET 58 58 M 1 (IN RES. 2).
 FT COMPLET 177 177 G -> R (IN RES. 3).
 SQ SEQUENCE 334 AA: 37547 MW: 36674.7043307AD98 GRC64;
 Query Match: 2.28, Score 9, DB 1, Length 334,
 Best Local Similarity: 100.0%, Ident. No. 0.42,
 Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 135 GSCWAFSA 143

RESULT 15
 CATL_PIG STANDARD: PRT: 334 AA.
 AC 028944:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cathepsin L precursor (NC 3.4.22.15).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae, Sus.
 OX NCBI_Taxid:9623;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Epididymis;
 RX MEDLINE:9606941; PubMed:7492581;
 RA Okamura N., Tanaka M., Uchiyama Y., Sugita Y., Endohs E.,
 RA Syntin P., Dacheux J.L.;
 RT "Direct evidence for the elevated synthesis and secretion of
 RT procathepsin L in the distal caput epididymis of boar."
 RL Biochim. Biophys. Acta 1245:221-226(1995)
 CC -1- FUNCTION: IMPORTANT FOR THE OVERALL MAINTENANCE OF EFFECTIVE IN
 CC LYSOSOMES.
 CC -1- CATALYTIC ACTIVITY: Specifically cleaves to that of pepsin. AS
 CC compared to pepsin B, cathepsin L exhibits higher activity
 CC towards protein substrates, but has little activity on 6 Arg Arg
 CC NIMEX, and no peptidyl-dipeptidase activity.
 CC -1- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE
 CC BONDS.
 CC -1- SUBCELLULAR LOCATION: lysosomal.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAEN FAMILY OF THIOL PROTEASES.
 CC
 CC This SWISS-Prot entry is derived through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement with EMBL. ***. (http://www.ebi.ac.uk/)
 CC or send an email to license@lsb.slb.ch).
 CC
 CC EMBL: D37917; BA07140.1;
 CC DR HSP: P07711; ICDL;
 CC DR MEMOPS: C01.032;
 CC DR InterPro: IPR000668; Peptidase_C1;
 CC DR InterPro: IPR000169; Thiolprol_act_site;
 CC DR Pfam: PF00112; Peptidase_C1; 1;
 CC DR PRINTS: PR00705; PAPAEN;
 CC DR PROSITE: PS00139; THIOL_PROTEASP_CYS1;
 CC DR PROSITE: PS00639; THIOL_PROTEASE_HIS1;
 CC DR PROSITE: PS00640; THIOL_PROTEASE_ASN1;
 CC DR HydroLase: Thiol protease, Glycyl-protein, lysosomal, zymogen; Signal;
 KM

FT SIGNAL 1 17 POTENTIAL.
 FT PROPEP 18 117 ACTIVATION PEPTIDE.
 FT CHAIN 118 289 CATHEPSIN L, HEAVY CHAIN.
 FT PROPEP 290 291 BY SIMILARITY.
 FT CHAIN 292 444 CATHEPSIN L, LIGHT CHAIN.
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT ACT_SITE 277 277 BY SIMILARITY.
 FT ACT_SITE 301 301 BY SIMILARITY.
 FT DISULFID 169 212 BY SIMILARITY.
 FT DISULFID 270 243 INTERCHAIN (BY SIMILARITY).
 FT CAPSHEID 222 222 N-LINKED (GLYCAN...)
 FT CAPSHEID 272 272 N-LINKED (GLYCAN...)
 SQ SEQUENCE 444 AA: 57176 MW: 57176.00000 GRC64;
 Query Match: 2.28, Score 9, DB 1, Length 444;
 Best Local Similarity: 100.0%, Ident. No. 0.42,
 Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 135 GSCWAFSA 143

RESULT 16
 CATL_RAT STANDARD: PRT: 444 AA.
 AC P07154; Q9QV07;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-SEP-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cathepsin L precursor (NC 3.4.22.15) (Cathepsin L) (CHP)
 DE (Cyclic protein-2) (CP-2).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Muridae; Rattus;
 OX NCBI_Taxid:10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STEIN-WISNIAK; TISSUE-Kidney;
 RX MEDLINE:8832042; PubMed:666144;
 RA Elleder K., Tokarz P., Elleder S., Kwasny J., Komlani E.,
 RA Katunuma N., Suzuki K.;
 RT "Molecular cloning and sequencing of cDNA for rat cathepsin L."
 RL Eur. J. Biochem. 225:73-76(1987).
 CC
 CC -1- FUNCTION: IMPORTANT FOR THE OVERALL MAINTENANCE OF EFFECTIVE IN
 CC LYSOSOMES.
 CC -1- CATALYTIC ACTIVITY: Specifically cleaves to that of pepsin. AS
 CC compared to pepsin B, cathepsin L exhibits higher activity
 CC towards protein substrates, but has little activity on 6 Arg Arg
 CC NIMEX, and no peptidyl-dipeptidase activity.
 CC -1- SUBUNIT: DIMER OF A HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE
 CC BONDS.
 CC -1- SUBCELLULAR LOCATION: lysosomal.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAEN FAMILY OF THIOL PROTEASES.
 CC
 CC This SWISS-Prot entry is derived through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement with EMBL. ***. (http://www.ebi.ac.uk/)
 CC or send an email to license@lsb.slb.ch).
 CC
 CC EMBL: D37917; BA07140.1;
 CC DR HSP: P07711; ICDL;
 CC DR MEMOPS: C01.032;
 CC DR InterPro: IPR000668; Peptidase_C1;
 CC DR InterPro: IPR000169; Thiolprol_act_site;
 CC DR Pfam: PF00112; Peptidase_C1; 1;
 CC DR PRINTS: PR00705; PAPAEN;
 CC DR PROSITE: PS00139; THIOL_PROTEASP_CYS1;
 CC DR PROSITE: PS00639; THIOL_PROTEASE_HIS1;
 CC DR PROSITE: PS00640; THIOL_PROTEASE_ASN1;
 CC DR HydroLase: Thiol protease, Glycyl-protein, lysosomal, zymogen; Signal;
 KM

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC PAPAEN FAMILY OF THIOL PROTEASES.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/infocentre/
CC or send an email to licencelists@ebi.ac.uk).

DR EMBL: Y14734: CAAT5029.1; -
DR EMBL: A000128: BAA25609.1; -
DR EMBL: A070448: AAT28906.1; -
DR EMBL: A019534: BAA3435.1; -
DR PDB: 1FH0: 30-JUL-01.
DR MIM: 603308; -
DR InterPro: IPR000668: Peptidase_c1;
DR InterPro: IPR000169: Thiolprol_act_site;
DR Pfam: PF00112: peptidase_c1; 1;
DR PRINTS: PR00705: PAPAEN.
DR PROSITE: PS00134: THIOL_PROTEASE_CYS; 1;
DR PROSITE: PS00649: THIOL_PROTEASE_HIS; 1;
DR PROSITE: PS00640: THIOL_PROTEASE_ASN; 1;
KW Hydrolyase; Thiol protease; glycoprotein; lysosomal; zymogen; signal;
KW 3D-structure; 1;
FT SIGNAL: 1 17
FT PROPEP: 18 113
FT CHAIN: 114 334
FT ACT_SITE: 138 138
FT ACT_SITE: 276 276
FT ACT_SITE: 301 301
FT DISULFID: 135 178
FT DISULFID: 169 211
FT DISULFID: 279 323
FT CARBOHYD: 221 221
FT CARBOHYD: 292 292
FT CONFLICT: 81 81
SQ SEQUENCE 334 AA: 37329 MW: 62055 IACSLQI2C8 CRG64;
Query Match 2.2% Score 9; DB 1; Length 234;
Fast Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSA 205
DB 135 CGSCWAFSA 143

RESULT 18
CYS1 LEIPI STANDARD: PRT: 354 AA.
ID CYS1 LEIPI
AC P35571;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Cysteine proteinase 1 precursor (EC 3.4.22.1) (Amaelgote cysteine
DE proteinase A-1)
GN CYS1.
OS Leishmania pifanoi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCST:taxid:5682;
RN [1]
RP SEQUENCE FROM N.A.
RA Almeida R.W.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases
FN [2]
PP SEQUENCE OF 147-311 FROM N.A.
PX MEDLINE:93149195; PubMed:8426065;
RA Traub-Cosco Y.M.; Dubois M.; Boukari L.F.; Mouton-Prat B.;
RT Identification of two distinct cysteine proteinase genes of
RT Leishmania pifanoi axenic amastigotes using the polymerase chain
RT reaction.*;

KL Mol. Biochem. Parasitol. 57:101-115(1993).
CC -1- FUNCTION: THE CYSTEINE PROTEINASE HAVE A PEPTIDON RELE IN
CC POST-PARASITIC INTERACTION AND VIRULENCE.
CC 1- DEVELOPMENTAL STAGE: PREVALENTLY EXPRESSED IN THE AMASTIGOTE STAGE.
CC EXPRESSED 4 TIMES MORE IN AMASTIGOTES THAN IN PROMASTIGOTES.
CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC PAPAEN FAMILY OF THIOL PROTEASES.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/infocentre/
CC or send an email to licencelists@ebi.ac.uk).

DR EMBL: L29149: AAG3909.1; -
DR EMBL: L09717: L_NCT_ABR-TALH_CDS.
DR EMBL: P25776: TALM.
DR MEROVS: C01.076; -
DR InterPro: IPR000668: peptidase_c1;
DR InterPro: IPR000169: Thiolprol_act_site;
DR Pfam: PF00112: peptidase_c1; 1;
DR PRINTS: PR00705: PAPAEN.
DR PROSITE: PS00134: THIOL_PROTEASE_CYS; 1;
DR PROSITE: PS00649: THIOL_PROTEASE_HIS; 1;
DR PROSITE: PS00640: THIOL_PROTEASE_ASN; 1;
KW Hydrolyase; Thiol protease; glycoprotein; zymogen; signal;
KW 3D-structure; 1;
FT SIGNAL: 1 24
FT PROPEP: 25 125
FT CHAIN: 126 374
FT ACT_SITE: 154 154
FT ACT_SITE: 289 289
FT ACT_SITE: 309 309
FT DISULFID: 150 191
FT DISULFID: 181 229
FT DISULFID: 282 332
FT CARBOHYD: 209 209
FT CONFLICT: 119 119
FT CONFLICT: 152 152
FT CONFLICT: 296 296
SQ SEQUENCE 374 AA: 38746 MW: 47550 IYDQDQAAA CRG34;
Query Match 2.2% Score 9; DB 1; Length 354;
Fast Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSA 205
DB 150 CGSCWAFSA 158

RESULT 19
LCPA LEIPI STANDARD: PRT: 454 AA.
ID LCPA LEIPI
AC P25776;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, last sequence update)
DT 01-MAY-1992 (Rel. 22, last annotation update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Cysteine proteinase A precursor (EC 3.4.22.1).
GN LCPA.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCST:taxid:5665;
RN [1]
RP SEQUENCE FROM N.A.
PP STRAIN L31A16; GHT: 93-13-3479;
PX MEDLINE:92274842; PubMed:1508041;
RA Mottlam J.C.; Robertson C.D.; Cosmids G.H.; Barry J.D.;
RT "A developmental stage revealed cysteine proteinase gene of Leishmania
RT mexicana".
KL Mol. Microbiol. 6:119-125 (1992/1992).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL LIFE CYCLE STAGES BUT AT

BA Watanabe H., Abe K., Emori Y., Hosoyama H., Arai S.,
 PT "Molecular cloning and gibberellin-induced expression of multiple
 RT cysteine proteinases of rice seeds (cytochrome)." J.
 RL J. Biol. Chem. 266:16897-16902(1991).
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN SEEDS.
 CC -1- INDUCTION: BY GIBBERELLIC ACID (GA).
 CC -1- STIMULATORY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAIN FAMILY OF THIOL PROTEASES.
 CC
 CC This SWISS-PROT entry is copyrighted. It is released through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, changed or used for commercial
 CC purposes. For more information, see: <http://www.ebi.ac.uk/seqdb/doc/doc.html>
 CC or send an email to license@ebi.ac.uk.
 CC
 CC PMID: 160406; NAI14402.1.
 DR PIR: J00388; KHR20A.
 DR HSP: P00745; ZACT.
 DR MEPS: C01.028; .
 DR InterPro: IPR000118; Granulin.
 DR InterPro: IPR000668; Peptidase_C1
 DR InterPro: IPR00169; Thiolprotease_act_site
 DR Pfam: PF00466; Granulin; 1
 DR Pfam: PF00112; Peptidase_C1; 1
 DR PRINTS: PG9705; PAPAIN.
 DR SMART: SM00277; GRAN_1.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 KW Hydrolyses thiol protease, cysteine, hydrolytic, signal.
 FT SIGNAL 1 21
 FT PROPEP 22 128
 FT CHAIN 129 458
 FT ACT_SITE 153 154
 FT ACT_SITE 289 289
 FT ACT_SITE 309 309
 FT DISULFID 150 192
 FT DISULFID 184 225
 FT DISULFID 284 344
 FT CARBOHYD 445 445
 SQ SEQUENCE 458 AA; 50274 MW; HIA090750735007 GP264;
 Query Match 2.2% Score 9; EB 1; Length 458,
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 9; Conservative 0; Missives 0; In-its 0; Type 0;
 QY 197 GSGTMAFSA 205
 DB 150 GSGTMAFSA 158
 RESULT 22
 CYSPTRYCR STANDARD; PRT; 467 AA.
 AC P25779;
 DT 01-MAY-1992 (Ref. 22, created)
 DT 01-MAY-1992 (Ref. 22, last sequence update)
 DT 16-OCT-2001 (Ref. 49, last annotation update)
 DE Crucipain protease (EC 3.4.22.1) (Major cysteine proteinase)
 DE (Crucipain).
 OS Trypanosoma cruzi.
 OC Eukaryota; Eukaryotes; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NBTAXID:5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TULAHUEN.
 RA MEDLINE:9218393; PubMed 155980;
 RA Parkin A.F., Mills A.A., Harth G., McKerrow J.H., Crank G.S.;
 RT "The sequence, organization and expression of the major cysteine
 RL protease (crucipain) from Trypanosoma cruzi." J. Biol. Chem. 267:7411-7420(1992).

EN [2]
 EN SEQUENCE FROM N.A. (Genbank accession 2 and 1800-4).
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 9218393; PubMed 141054;
 RA Campetella C., Fontanesa J., Masferrer L., Trush A.G.;
 RA McKerrow J.H., Crank G.S.;
 RA "The major cysteine proteinase (crucipain) from Trypanosoma cruzi is
 RT encoded by multiple polymorphic tandemly organized genes located on
 RT different chromosomes." J.
 RL Mol. Biochem. Parasitol. 50:225-234(1992).
 EN [3]
 EN SEQUENCE OF 141 kDa FROM N.A.
 RC STRAIN KAT.
 RC MEDLINE 901999; PubMed 2406590;
 RA Parkin A.F., Trush A.G., McKerrow J.H.;
 RA "Amplification and sequencing of genomic DNA fragments encoding
 RT cysteine proteinases from protozoan parasites." J.
 RL Mol. Biochem. Parasitol. 39:1-8(1990).
 EN [4]
 EN SEQUENCE OF 205 kDa FROM N.A.
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 9124625; PubMed 2038564;
 RA Masferrer L., Fontanesa J., Campetella C., Trush A.G.;
 RA McKerrow J.H., Crank G.S.;
 RA "The C-terminal extension of the major cysteine proteinase
 RT (crucipain) from Trypanosoma cruzi." J.
 RL Mol. Biochem. Parasitol. 49:47-56(1991).
 EN [5]
 EN AUTOCATALYSIS OF C-TERMINAL.
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 9118294; PubMed 201151;
 RA Bellan M., Watanabe H., Crank G.S.;
 RA "Self-proteolysis of the cysteine proteinase, crucipain, from
 RT Trypanosoma cruzi gives a major fragment corresponding to the
 RT carboxy-terminal domain." J.
 RL Mol. Biochem. Parasitol. 44:15-21(1991).
 EN [6]
 EN SPECIFICITY.
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 9810712; PubMed 2407295;
 RA Garcia J.J., Garcia-Franco M.C., Martinez J.;
 RA "Purification and characterization of a cysteine proteinase (crucipain) from
 RT Trypanosoma cruzi." J.
 RL Biochim. Biophys. Acta 1037:186-191(1990).
 EN [7]
 EN SEQUENCE OF 127 146 AND 304-417.
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 8921914; PubMed 2561942;
 RA Garcia J.J., Garcia-Franco M.C., Martinez J.;
 RA "Purification and characterization of a cysteine proteinase (crucipain) from
 RT Trypanosoma cruzi." J.
 RL Mol. Biochem. Parasitol. 44:15-21(1991).
 EN [8]
 EN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 121-147.
 RC STRAIN-TULAHUEN 2;
 RA MEDLINE 9740690; PubMed 9260273;
 RA Gilmer S.A., Crank G.S., Fontanesa J.;
 RA "Structural determinants of specificity in the cysteine proteinase
 RT crucipain." J.
 RL Protein Sci. 6:1603-1611(1997).
 CC 1. FUNCTION: INTERACTS WITH C-TERMINAL PEPTIDES AND WITH THE
 CC C-TERMINAL PEPTIDE OF THE PARASITIC ANTIMETABOLITE, AND WITH
 CC VALORIN, WITHIN THE TERMINAL AND C-TERMINAL AND THE ACTIVE SITES
 CC GROUP.
 CC 2. FURTHER INFORMATION: THE C-TERMINAL PEPTIDE MAY PLAY AN IMPORTANT ROLE IN THE
 CC DEVELOPMENT AND DIFFERENTIATION OF THE PARASITIC ANTIMETABOLITE, AND
 CC OF THEIR LIFE CYCLE.
 CC 3. ENZYME REGULATION: SPECIFICALLY INHIBITED BY 1,4-ALPHA-AMINO
 CC BROMOSUCCINYLAMIDES (4-GUANIDINO-2-PYRIDYL) (MEDLINE 9740690, AND
 CC N-ALPHA-P-TREYLA-L-LEUCINE CHLOROMETHYL KETONE).
 CC 4. DEVELOPMENTAL STAGE: PRESENT IN ALL DEVELOPMENTAL STAGES.
 CC 5. MISCELLANEOUS: FORMED BY CRUCIPAIN IS ABLE TO DEGRADE ITSELF,
 CC YIELDING A COMPLEX MIXTURE OF SMALL PEPTIDES, AND A WATER-SOLUBLE
 CC FRAGMENT.

| FW | Hydrophobic | Thiol | Proximal | Glycyl-proline | Lysosomal | Zymogenic | Stimulated |
|---------------|-------------|--------------|--------------|------------------|--------------------|-----------|------------|
| FF | SIGNAL | 1 | 2 | POTENTIAL | | | |
| FF | PROPEL | 2 | 104 | ACTIVATION | FEELING (FOOTBALL) | | |
| FF | CHAIN | 105 | 419 | CATHETER | 1 | | |
| FF | ACT SITE | 123 | 129 | BY SIMILARITY | | | |
| FF | ACT SITE | 127 | 275 | BY SIMILARITY | | | |
| FF | ACT SITE | 286 | 286 | BY SIMILARITY | | | |
| FF | DISULFID | 126 | 107 | BY SIMILARITY | | | |
| FF | DISULFID | 160 | 230 | BY SIMILARITY | | | |
| FF | DISULFID | 238 | 307 | BY SIMILARITY | | | |
| SS | SEQUENCE | 419 AA: | 46186 MW: | 90F55A34E1194873 | CR7641 | | |
| Quality Match | | | | | | | |
| Host Local | Similarity | 100.0% | Score: 82.18 | Length: 419 | | | |
| Matches | 97 | Conservation | 07 | Mutations | 47 | Indels | 07 |
| 07 | 197 | CONSWARE | 204 | | | | |
| | | | | | | | |
| DB | 126 | CONSWARE | 133 | | | | |

| | | | | |
|------------|---|-----------|------|---------|
| MMAL-DEKPT | | STANDARD: | PRI: | 320 AA. |
| ID | MMAL-DEKPT | | | |
| AC | P08176; Q24116; | | | |
| D1 | 01-AUG-1988 (Rel. 08) (Created) | | | |
| D2 | 01-FEB-1995 (Rel. 41) (Last sequence update) | | | |
| D3 | 16-OCT-2001 (Rel. 40) (Last annotation update) | | | |
| D4 | Major mite fecal antigen Der p 1 precursor (EC 3.4.22.) (Der p 1)-
DERP1. | | | |
| D5 | Intermolecular photorecognition (House-dust mite). | | | |
| OR | Eukaryotic Metazoa: Arthropoda: Chelicerata: Arachnida: Acari:
Eriophora: Mesostigmata: Ixodidae, Amblyomidae, Pyronotiphididae,
Trombidiformes. | | | |
| OX | Enviropedia:1458. | | | |
| RN | Nucleotide(s):6956; | | | |
| RX | 111 | | | |
| SC | SEQUENCE FROM N.A.S. AND POLYMERPHISM. | | | |
| RX | MEDLINE 934576BZ; PubMed 8354459; | | | |
| RA | Chua K.Y., Kohal P.K.; Thomas W.R.: | | | |
| RT | "Sequence polymorphisms of cDNA clones encoding the mite allergen Der
p 1." | | | |
| RL | Int. Arch. Allergy Immunol. 101:464-468(1993). | | | |
| RN | 121 | | | |
| RP | SEQUENCED AT 76,320 FROM N.A.S. | | | |
| RX | MEDLINE 8809541J; PubMed 3245840; | | | |
| RA | Chua K.Y., Stewart G.A.; Thomas W.R.; Simpson R.J.; Pilsworth T.M.;
Pilczwa T.M.; Turner K.O.: | | | |
| RT | "Sequence analysis of cDNA coding for a major house dust mite
allergen, Der p 1, homologized with cysteine proteases." | | | |
| RL | J. Exp. Med. 197:175-182(1988). | | | |
| RN | 131 | | | |
| RP | SEQUENCED AT 81,176 FROM N.A.S. | | | |
| RX | MEDLINE 861149D; PubMed 3276429; | | | |
| RA | Thomas W.R.; Stewart G.A.; Simpson R.J.; Chua K.Y.; Pilczwa T.M.;
O'Donoghue R.; Nickol A.; Turner K.O.: | | | |
| RT | "Cloning and expression of cDNA coding for the major house dust mite
allergen Der p 1 in Escherichia coli." | | | |
| RL | Int. Arch. Allergy Appl. Immunol. 95:127-129(1988). | | | |
| RN | 141 | | | |
| SC | SEQUENCE OF 99-127 | | | |
| SV | SEPTIME 86-09-09; PubMed 1372999; | | | |
| RA | Lind P.; Hultsen O.C.; Horn N.: | | | |
| RT | "The binding of mouse hyaluronidase and human IgE antibodies to the major
fecal allergen, Der p 1, of Dermatophagoides pteronyssinus, relative to
binding site location and species specificity studied by solid phase
inhibition assays with radiolabeled antigen." | | | |
| RL | J. Immunol. 140:4256-4262(1988). | | | |
| RN | 151 | | | |
| RV | KEYWORDS: 76 232 241 | | | |
| RX | MEDLINE 91215493; PubMed 2021874; | | | |
| RA | Dilworth R.J.; Chua K.Y.; Thomas W.R.: | | | |
| RT | "Sequence analysis of cDNA coding for a major house dust mite
allergen, Der p 1." | | | |

[illegible][illegible]

Query Match 2.98; Score 8; DB 1; Length 321;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GSCSWAFS 204
 |||||
 Db 130 GSCSWAFS 137

RESULT 33
 ID MMAL_DERFA STANDARD; PRT: 321 AA.

AC P16311;
 DT 01-AMC-1990 (Ref. 15, Created)
 DT 01-FEB-1995 (Ref. 31, Last sequence update)
 DT 16-OCT-2001 (Ref. 40, Last annotation update)
 DE Major mite fecal allergen Der F 1 precursor (Ref. 3, 4, 22, ...) (Der F 1).
 GN DERF1.
 OS Dermatophagoides farinæ (House-dust mite).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Acarina; Acari;
 OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
 OC Dermatophagoides.
 OX NCBI_TaxID=9554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91215492; PubMed=2021974;
 FA Blumworth P.J., Chua R.Y., Thomas W.E.;
 RT "Sequence analysis of cDNA coding for a major house dust mite
 RT allergen, Der f 1.";
 RL Clin. Exp. Allergy 21:25-32(1991).
 PN [2]
 RP SEQUENCE OF 98-309 FROM N.A.
 RA Kent N., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases
 RN [3]
 RP SEQUENCE OF 98-128
 RX MEDLINE=8279148; PubMed=3372999;
 RA Lind P., Hansen O.C., Horn N.;
 RT "The binding of mouse hyaluronidase and human IgE antibodies to the major
 RT fecal allergen, Der f 1, of dermatophagoides pteronyssinus: Relation
 RT binding site location and species specificity studied by solid phase
 RT inhibition assays with radiolabeled antigen.";
 RL J. Immunol. 149:4256-4262(1988).
 CC -1- FUNCTION: THIS PROTEIN IS THE MAJOR ALLERGEN OF HOUSE DUST MITE,
 CC IT IS A THIOL PROTEASE THAT HYDROLYSES PROTEINS, WITH A PREFERENCE
 CC FOR PHE OR BASIC RESIDUES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAIN FAMILY OF THIOL PROTEASES.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, contribution -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/doc/usage>
 CC or send an email to license@isb.ch).

EMBL: X65196; CAA6316.1;
 DR PIR: A27634; A27634.
 DR HSSP: O17473; 1BMO.
 DR MEROPS: C01.073;
 DR InterPro: IPR000169; Thiolprotease_act.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00649; THIOL_PROTEASE_HIS; 1.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 KW Hydrolase; Thiol protease; Allergen; Cysteine, Cysteine, Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 98 ACTIVATION PEPTIDE.

FT CHAIN 99 321
 FT ACT_SITE 133 133
 FT ACT_SITE 269 269
 FT ACT_SITE 288 288
 FT CATALYTIC 151 151
 FT DISULFID 162 216
 FT DISULFID 170 170
 FT DISULFID 164 202
 FT COMPLET 201 201
 FT COMPLET 282 282
 SO SEQUENCE 321 AA; 364 kD; MW: 0452454568476476; CMOCA;

Query Match 2.98; Score 8; DB 1; Length 321;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GSCSWAFS 204
 |||||
 Db 130 GSCSWAFS 137

RESULT 34
 ID CYSL_HOMAM STANDARD; PRT: 322 AA.

AC P13272;
 DT 01-AMC-1990 (Ref. 13, Created)
 DT 01-MAY-1992 (Ref. 22, Last sequence update)
 DT 01-NOV-1995 (Ref. 22, Last annotation update)
 DE Digestive cysteine proteinase 1 precursor (Ref. 3, 4, 22, ...).
 GN LGP1.
 OS Homarus americanus (American lobster).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Malacostraca;
 OC Eumalacostraca; Decapoda; Decapoda; Eumalacostraca; Astacodea;
 OC Neopropodea; Neopropodea; Homarus.
 OX NCBI_TaxID=6706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92070467; PubMed=1659690;
 FA Laycock M.V., Murray R.M., di Francesco M., Gallant J.W.;
 RT "Molecular cloning of three cDNAs that encode cysteine proteinases in
 RT the digestive gland of the American lobster (Homarus americanus).";
 RL FEBS Lett. 292:115-120(1991).
 PN [2]
 RP SEQUENCE OF 106-133
 KC TISSUE: Digestive juices;
 RX MEDLINE=9030876; PubMed=2597115;
 FA Laycock M.V., Murray R.M., di Francesco M., Storer A.C.;
 RT "Purification and characterization of a digestive cysteine proteinase
 RT from the American lobster (Homarus americanus).";
 RL Biochem. J. 263:439-444(1989).
 CC -1- SERINE ACTIVATION: INHIBITED BY E-64, EDTA, L-PROLINE, L-PROLINE, HEPES,
 CC METAL IONS, TOSAMIC ACID, DITHIONITROBENZENE, P-HYDROXYMERKURY-
 CC MERCAPTO, ACTIVATED BY MERCAPTOETHANOL AND DITHIONITROBENZENE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAIN FAMILY OF THIOL PROTEASES.
 CC -1- SIMILARITY: BELONGS TO PAPAIN AND CATHEPSIN L.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, contribution -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/doc/usage>
 CC or send an email to license@isb.ch).

EMBL: X63567; CAA5127.1;
 DR PIR: S06154; S06154.
 DR PIR: S19649; S19649.
 DR HSSP: P07711; 1CJ1.
 DR MEROPS: C01.040;
 DR InterPro: IPR000169; Thiolprotease_act.
 DR InterPro: IPR000169; Thiolprotease_act.

[illegible]


```

RT      "Identification and characterization of the Cydia pomonella
RT      granulin virus cathpsin and chitinase genes."
RT      J Gen Virol 79:2283-2292(1998)
P1      - FUNCTION: MAY PLAY A ROLE IN DEGRADATION OF INFECTED LARVAE TO
CC      FACILITATE HORIZONTAL TRANSMISSION OF THE VIRUS.
CC      - SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC      PAPAIN FAMILY OF THIOL PROTEASES.
CC      -----
CC      CC This SWISS-Prot entry is copyright © 1996 by the Swiss Institute
CC      CC for Bioinformatics. All rights reserved. No part of this
CC      CC entry may be reproduced or transmitted in any form or by any
CC      CC means electronic or mechanical, including photocopying, recording,
CC      CC or by information storage and retrieval systems, without prior
CC      CC written permission from the publisher, Birkbeck College, University
CC      CC of London. See also the following URL: http://www.birkbeck.ac.uk/~bioc/
CC      CC
CC      EMBL: AB010886; BAA31201.1; ...
DR      HSSP: P25776; IATM
DR      MEROPS: G01.083; ...
DR      InterPro: IPRO00668; Peptidase_C1
DR      InterPro: IPRO00169; Thiolpro_act_site
DR      Pfam: PF00112; Peptidase_C1_1
DR      PRINTS: PR00705; PAPAIN.
DR      ProSITE: PS00154; THIOL_PROTEASE_ASYT_1
DR      ProSITE: PS00640; THIOL_PROTEASE_HIS_1
DR      ProSITE: PS00640; THIOL_PROTEASE_ASN_1
KW      Hydrolase; Thiol protease.
FT      ACT_SITE 148..148 BY SIMILARITY.
FT      ACT_SITE 289..289 BY SIMILARITY.
FT      ACT_SITE 300..300 BY SIMILARITY.
FT      DISULETD 145..186 BY SIMILARITY.
FT      DISULFID 179..219 BY SIMILARITY.
FT      DIELFID 379..379 BY SIMILARITY.
SO      SEQUENCE 333 AA, 37433 MW, 7510C9AC3AHTF5A CRC64;

Query Match          2.0% Score 8; DB 1; Length 333,
Best Local Similarity 100.0%; Fwd. No. 3; r
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      197 CGSWAFS 204
DB      145 CGSWAFS 152

RESULT 44
CATL_DROME
ID CATL_DROME STANDARD: PROT: 341 AA.
AC Q95025; O97431;
DT 1-oct-2001 (Seq) 40; (reated)
DT 16-oct-2001 (Bel) 40; Last sequence update)
DI 16-oct-2001 (bel) 40; Last annotation update)
DF Cathpsin 1 proteinase (EC 3.4.22.15) (cysteine protease 1),
GN Cpl OR FS(250CA OR CG5692).
OS Drosophila melanogaster (Drosophila)
PE Pikaforia; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Phyllophaga; Hymenoptera; Diptera; Brachycera; Muscidae;
OC Phoridae; Drosophilidae; Drosophila
OX NCBI_TaxID=7227;
RN RN
RP RP
RP SEQUENCE FROM N.A.
RX MEHLINE_97254146; PubMed=9099981;
RX MEHLINE_97254146; PubMed=9099981;
RA Tryptosus ty. Baltimore IV.
RE Tryptosus ty. Baltimore IV.
RT The Drosophila melanogaster haemocyte cell line mbo-2."
RL Insect Mol Biol 6:173-181(1997).
[2]
SEQUENCE FROM N.A.
RX MEHLINE_98324856; PubMed=9662477;
RX Gray Y H M Seng J A.; Protein C R , Enghols W R ;
RT "Structure and associated mutational effects of the cysteine
RT proteinase (Cpl) gene of Drosophila melanogaster."
RL Insect Mol Biol 7:291-294(1998)

```

[illegible]



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuLink Ltd.

OM protein - protein search, using sw model

Run : 10 Search 9, 2002, 11:46:13, Search time 13 seconds
(without alignments)
2992102 Million cell updates/sec

Title: US-09-598-062-4

Sequence: 1 MDISNNVPHQRTSPYIA... KRRTGGLVPPVYVINGST 401

Scoring table: OLIGO
Gapop 60 0, Gapext 60 0

Searched: 562222 seqs, 1729499 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: listing first 1000 matches

Database: SPTPMM1_19.*

- 1: SP_archaea.*
- 2: SP_bacteria.*
- 3: SP_fungi.*
- 4: SP_human.*
- 5: SP_invertebrate.*
- 6: SP_mammal.*
- 7: SP_mhc.*
- 8: SP_organelle.*
- 9: SP_phage.*
- 10: SP_plant.*
- 11: SP_protist.*
- 12: SP_virus.*
- 13: SP_vertebrate.*
- 14: SP_unclassified.*
- 15: SP_virus.*
- 16: SP_bacteriap.*
- 17: SP_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

| Result No | Score | Query Match | Length DB ID | Description |
|-----------|-------|-------------|--------------|-------------|
| 1 | 401 | 100.0 | 401 5 | Q9Y174 |
| 2 | 112 | 3.8 | 344 10 | Q9Y206 |
| 3 | 12 | 3.0 | 466 10 | Q49877 |
| 4 | 12 | 4.6 | 162 12 | P89206 |
| 5 | 11 | 2.7 | 301 6 | Q15566 |
| 6 | 11 | 2.7 | 462 10 | Q91727 |
| 7 | 11 | 2.7 | 462 10 | Q42774 |
| 8 | 11 | 2.7 | 484 5 | Q98834 |
| 9 | 11 | 2.7 | 484 5 | Q9N842 |
| 10 | 11 | 2.7 | 484 5 | Q9N842 |
| 11 | 11 | 2.7 | 484 5 | Q9N842 |
| 12 | 11 | 2.7 | 484 5 | Q9N842 |
| 13 | 10 | 2.5 | 217 5 | Q9N842 |
| 14 | 10 | 2.5 | 217 5 | Q9N842 |
| 15 | 10 | 2.5 | 217 5 | Q9N842 |
| 16 | 10 | 2.5 | 217 5 | Q9N842 |

| | | | | |
|-----|----|-----|-------|--------|
| 17 | 10 | 2.5 | 326 5 | Q95VA7 |
| 18 | 10 | 2.5 | 381 5 | Q96GN6 |
| 19 | 10 | 2.5 | 394 5 | Q95NF6 |
| 20 | 10 | 2.5 | 443 5 | P92185 |
| 21 | 10 | 2.5 | 443 5 | Q18698 |
| 22 | 10 | 2.5 | 443 5 | P96528 |
| 23 | 10 | 2.5 | 443 5 | Q95M87 |
| 24 | 10 | 2.5 | 443 5 | Q92202 |
| 25 | 9 | 2.2 | 144 5 | Q96N83 |
| 26 | 9 | 2.2 | 144 5 | Q9N842 |
| 27 | 9 | 2.2 | 144 5 | Q9N842 |
| 28 | 9 | 2.2 | 144 5 | Q9N842 |
| 29 | 9 | 2.2 | 144 5 | Q9N842 |
| 30 | 9 | 2.2 | 144 5 | Q9N842 |
| 31 | 9 | 2.2 | 144 5 | Q9N842 |
| 32 | 9 | 2.2 | 144 5 | Q9N842 |
| 33 | 9 | 2.2 | 144 5 | Q9N842 |
| 34 | 9 | 2.2 | 144 5 | Q9N842 |
| 35 | 9 | 2.2 | 144 5 | Q9N842 |
| 36 | 9 | 2.2 | 144 5 | Q9N842 |
| 37 | 9 | 2.2 | 144 5 | Q9N842 |
| 38 | 9 | 2.2 | 144 5 | Q9N842 |
| 39 | 9 | 2.2 | 144 5 | Q9N842 |
| 40 | 9 | 2.2 | 144 5 | Q9N842 |
| 41 | 9 | 2.2 | 144 5 | Q9N842 |
| 42 | 9 | 2.2 | 144 5 | Q9N842 |
| 43 | 9 | 2.2 | 144 5 | Q9N842 |
| 44 | 9 | 2.2 | 144 5 | Q9N842 |
| 45 | 9 | 2.2 | 144 5 | Q9N842 |
| 46 | 9 | 2.2 | 144 5 | Q9N842 |
| 47 | 9 | 2.2 | 144 5 | Q9N842 |
| 48 | 9 | 2.2 | 144 5 | Q9N842 |
| 49 | 9 | 2.2 | 144 5 | Q9N842 |
| 50 | 9 | 2.2 | 144 5 | Q9N842 |
| 51 | 9 | 2.2 | 144 5 | Q9N842 |
| 52 | 9 | 2.2 | 144 5 | Q9N842 |
| 53 | 9 | 2.2 | 144 5 | Q9N842 |
| 54 | 9 | 2.2 | 144 5 | Q9N842 |
| 55 | 9 | 2.2 | 144 5 | Q9N842 |
| 56 | 9 | 2.2 | 144 5 | Q9N842 |
| 57 | 9 | 2.2 | 144 5 | Q9N842 |
| 58 | 9 | 2.2 | 144 5 | Q9N842 |
| 59 | 9 | 2.2 | 144 5 | Q9N842 |
| 60 | 9 | 2.2 | 144 5 | Q9N842 |
| 61 | 9 | 2.2 | 144 5 | Q9N842 |
| 62 | 9 | 2.2 | 144 5 | Q9N842 |
| 63 | 9 | 2.2 | 144 5 | Q9N842 |
| 64 | 9 | 2.2 | 144 5 | Q9N842 |
| 65 | 9 | 2.2 | 144 5 | Q9N842 |
| 66 | 9 | 2.2 | 144 5 | Q9N842 |
| 67 | 9 | 2.2 | 144 5 | Q9N842 |
| 68 | 9 | 2.2 | 144 5 | Q9N842 |
| 69 | 9 | 2.2 | 144 5 | Q9N842 |
| 70 | 9 | 2.2 | 144 5 | Q9N842 |
| 71 | 9 | 2.2 | 144 5 | Q9N842 |
| 72 | 9 | 2.2 | 144 5 | Q9N842 |
| 73 | 9 | 2.2 | 144 5 | Q9N842 |
| 74 | 9 | 2.2 | 144 5 | Q9N842 |
| 75 | 9 | 2.2 | 144 5 | Q9N842 |
| 76 | 9 | 2.2 | 144 5 | Q9N842 |
| 77 | 9 | 2.2 | 144 5 | Q9N842 |
| 78 | 9 | 2.2 | 144 5 | Q9N842 |
| 79 | 9 | 2.2 | 144 5 | Q9N842 |
| 80 | 9 | 2.2 | 144 5 | Q9N842 |
| 81 | 9 | 2.2 | 144 5 | Q9N842 |
| 82 | 9 | 2.2 | 144 5 | Q9N842 |
| 83 | 9 | 2.2 | 144 5 | Q9N842 |
| 84 | 9 | 2.2 | 144 5 | Q9N842 |
| 85 | 9 | 2.2 | 144 5 | Q9N842 |
| 86 | 9 | 2.2 | 144 5 | Q9N842 |
| 87 | 9 | 2.2 | 144 5 | Q9N842 |
| 88 | 9 | 2.2 | 144 5 | Q9N842 |
| 89 | 9 | 2.2 | 144 5 | Q9N842 |
| 90 | 9 | 2.2 | 144 5 | Q9N842 |
| 91 | 9 | 2.2 | 144 5 | Q9N842 |
| 92 | 9 | 2.2 | 144 5 | Q9N842 |
| 93 | 9 | 2.2 | 144 5 | Q9N842 |
| 94 | 9 | 2.2 | 144 5 | Q9N842 |
| 95 | 9 | 2.2 | 144 5 | Q9N842 |
| 96 | 9 | 2.2 | 144 5 | Q9N842 |
| 97 | 9 | 2.2 | 144 5 | Q9N842 |
| 98 | 9 | 2.2 | 144 5 | Q9N842 |
| 99 | 9 | 2.2 | 144 5 | Q9N842 |
| 100 | 9 | 2.2 | 144 5 | Q9N842 |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|--------|-------------------|-----|---|-----|-----|---|--------|-------------------------|
| 236 | 7 | 1.7 | 190 | 2 | Q91NM9 | Q91NM9 unclutered | 409 | 7 | 1.7 | 441 | 5 | Q25028 | Q25028 haemochromatosis |
| 237 | 7 | 1.7 | 197 | 2 | Q91NM9 | Q91NM9 unclutered | 410 | 7 | 1.7 | 442 | 5 | Q25031 | Q25031 haemochromatosis |
| 238 | 7 | 1.7 | 200 | 6 | Q91NM9 | Q91NM9 unclutered | 411 | 7 | 1.7 | 443 | 5 | Q25034 | Q25034 haemochromatosis |
| 239 | 7 | 1.7 | 206 | 10 | Q91NM9 | Q91NM9 unclutered | 412 | 7 | 1.7 | 444 | 5 | Q25037 | Q25037 haemochromatosis |
| 240 | 7 | 1.7 | 208 | 2 | Q91NM9 | Q91NM9 unclutered | 413 | 7 | 1.7 | 445 | 5 | Q25040 | Q25040 haemochromatosis |
| 241 | 7 | 1.7 | 209 | 2 | Q91NM9 | Q91NM9 unclutered | 414 | 7 | 1.7 | 446 | 5 | Q25043 | Q25043 haemochromatosis |
| 242 | 7 | 1.7 | 209 | 2 | Q91NM9 | Q91NM9 unclutered | 415 | 7 | 1.7 | 447 | 5 | Q25046 | Q25046 haemochromatosis |
| 243 | 7 | 1.7 | 219 | 10 | Q91NM9 | Q91NM9 unclutered | 416 | 7 | 1.7 | 448 | 5 | Q25049 | Q25049 haemochromatosis |
| 244 | 7 | 1.7 | 221 | 5 | Q91NM9 | Q91NM9 unclutered | 417 | 7 | 1.7 | 449 | 5 | Q25052 | Q25052 haemochromatosis |
| 245 | 7 | 1.7 | 226 | 5 | Q91NM9 | Q91NM9 unclutered | 418 | 7 | 1.7 | 450 | 5 | Q25055 | Q25055 haemochromatosis |
| 246 | 7 | 1.7 | 233 | 5 | Q91NM9 | Q91NM9 unclutered | 419 | 7 | 1.7 | 451 | 5 | Q25058 | Q25058 haemochromatosis |
| 247 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 420 | 7 | 1.7 | 452 | 5 | Q25061 | Q25061 haemochromatosis |
| 248 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 421 | 7 | 1.7 | 453 | 5 | Q25064 | Q25064 haemochromatosis |
| 249 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 422 | 7 | 1.7 | 454 | 5 | Q25067 | Q25067 haemochromatosis |
| 250 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 423 | 7 | 1.7 | 455 | 5 | Q25070 | Q25070 haemochromatosis |
| 251 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 424 | 7 | 1.7 | 456 | 5 | Q25073 | Q25073 haemochromatosis |
| 252 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 425 | 7 | 1.7 | 457 | 5 | Q25076 | Q25076 haemochromatosis |
| 253 | 7 | 1.7 | 233 | 12 | Q91NM9 | Q91NM9 unclutered | 426 | 7 | 1.7 | 458 | 5 | Q25079 | Q25079 haemochromatosis |
| 254 | 7 | 1.7 | 269 | 5 | Q91NM9 | Q91NM9 unclutered | 427 | 7 | 1.7 | 459 | 5 | Q25082 | Q25082 haemochromatosis |
| 255 | 7 | 1.7 | 269 | 11 | Q91NM9 | Q91NM9 unclutered | 428 | 7 | 1.7 | 460 | 5 | Q25085 | Q25085 haemochromatosis |
| 256 | 7 | 1.7 | 278 | 2 | Q91NM9 | Q91NM9 unclutered | 429 | 7 | 1.7 | 461 | 5 | Q25088 | Q25088 haemochromatosis |
| 257 | 7 | 1.7 | 281 | 6 | Q91NM9 | Q91NM9 unclutered | 430 | 7 | 1.7 | 462 | 5 | Q25091 | Q25091 haemochromatosis |
| 258 | 7 | 1.7 | 282 | 6 | Q91NM9 | Q91NM9 unclutered | 431 | 7 | 1.7 | 463 | 5 | Q25094 | Q25094 haemochromatosis |
| 259 | 7 | 1.7 | 286 | 11 | Q91NM9 | Q91NM9 unclutered | 432 | 7 | 1.7 | 464 | 5 | Q25097 | Q25097 haemochromatosis |
| 260 | 7 | 1.7 | 289 | 16 | Q91NM9 | Q91NM9 unclutered | 433 | 7 | 1.7 | 465 | 5 | Q25100 | Q25100 haemochromatosis |
| 261 | 7 | 1.7 | 293 | 2 | Q91NM9 | Q91NM9 unclutered | 434 | 7 | 1.7 | 466 | 5 | Q25103 | Q25103 haemochromatosis |
| 262 | 7 | 1.7 | 301 | 16 | Q91NM9 | Q91NM9 unclutered | 435 | 7 | 1.7 | 467 | 5 | Q25106 | Q25106 haemochromatosis |
| 263 | 7 | 1.7 | 303 | 6 | Q91NM9 | Q91NM9 unclutered | 436 | 7 | 1.7 | 468 | 5 | Q25109 | Q25109 haemochromatosis |
| 264 | 7 | 1.7 | 303 | 13 | Q91NM9 | Q91NM9 unclutered | 437 | 7 | 1.7 | 469 | 5 | Q25112 | Q25112 haemochromatosis |
| 265 | 7 | 1.7 | 305 | 4 | Q91NM9 | Q91NM9 unclutered | 438 | 7 | 1.7 | 470 | 5 | Q25115 | Q25115 haemochromatosis |
| 266 | 7 | 1.7 | 305 | 10 | Q91NM9 | Q91NM9 unclutered | 439 | 7 | 1.7 | 471 | 5 | Q25118 | Q25118 haemochromatosis |
| 267 | 7 | 1.7 | 305 | 11 | Q91NM9 | Q91NM9 unclutered | 440 | 7 | 1.7 | 472 | 5 | Q25121 | Q25121 haemochromatosis |
| 268 | 7 | 1.7 | 306 | 5 | Q91NM9 | Q91NM9 unclutered | 441 | 7 | 1.7 | 473 | 5 | Q25124 | Q25124 haemochromatosis |
| 269 | 7 | 1.7 | 306 | 11 | Q91NM9 | Q91NM9 unclutered | 442 | 7 | 1.7 | 474 | 5 | Q25127 | Q25127 haemochromatosis |
| 270 | 7 | 1.7 | 306 | 11 | Q91NM9 | Q91NM9 unclutered | 443 | 7 | 1.7 | 475 | 5 | Q25130 | Q25130 haemochromatosis |
| 271 | 7 | 1.7 | 307 | 11 | Q91NM9 | Q91NM9 unclutered | 444 | 7 | 1.7 | 476 | 5 | Q25133 | Q25133 haemochromatosis |
| 272 | 7 | 1.7 | 309 | 5 | Q91NM9 | Q91NM9 unclutered | 445 | 7 | 1.7 | 477 | 5 | Q25136 | Q25136 haemochromatosis |
| 273 | 7 | 1.7 | 314 | 5 | Q91NM9 | Q91NM9 unclutered | 446 | 7 | 1.7 | 478 | 5 | Q25139 | Q25139 haemochromatosis |
| 274 | 7 | 1.7 | 317 | 5 | Q91NM9 | Q91NM9 unclutered | 447 | 7 | 1.7 | 479 | 5 | Q25142 | Q25142 haemochromatosis |
| 275 | 7 | 1.7 | 322 | 5 | Q91NM9 | Q91NM9 unclutered | 448 | 7 | 1.7 | 480 | 5 | Q25145 | Q25145 haemochromatosis |
| 276 | 7 | 1.7 | 323 | 5 | Q91NM9 | Q91NM9 unclutered | 449 | 7 | 1.7 | 481 | 5 | Q25148 | Q25148 haemochromatosis |
| 277 | 7 | 1.7 | 324 | 10 | Q91NM9 | Q91NM9 unclutered | 450 | 7 | 1.7 | 482 | 5 | Q25151 | Q25151 haemochromatosis |
| 278 | 7 | 1.7 | 325 | 5 | Q91NM9 | Q91NM9 unclutered | 451 | 7 | 1.7 | 483 | 5 | Q25154 | Q25154 haemochromatosis |
| 279 | 7 | 1.7 | 329 | 5 | Q91NM9 | Q91NM9 unclutered | 452 | 7 | 1.7 | 484 | 5 | Q25157 | Q25157 haemochromatosis |
| 280 | 7 | 1.7 | 329 | 11 | Q91NM9 | Q91NM9 unclutered | 453 | 7 | 1.7 | 485 | 5 | Q25160 | Q25160 haemochromatosis |
| 281 | 7 | 1.7 | 329 | 11 | Q91NM9 | Q91NM9 unclutered | 454 | 7 | 1.7 | 486 | 5 | Q25163 | Q25163 haemochromatosis |
| 282 | 7 | 1.7 | 329 | 13 | Q91NM9 | Q91NM9 unclutered | 455 | 7 | 1.7 | 487 | 5 | Q25166 | Q25166 haemochromatosis |
| 283 | 7 | 1.7 | 331 | 4 | Q91NM9 | Q91NM9 unclutered | 456 | 7 | 1.7 | 488 | 5 | Q25169 | Q25169 haemochromatosis |
| 284 | 7 | 1.7 | 331 | 5 | Q91NM9 | Q91NM9 unclutered | 457 | 7 | 1.7 | 489 | 5 | Q25172 | Q25172 haemochromatosis |
| 285 | 7 | 1.7 | 331 | 5 | Q91NM9 | Q91NM9 unclutered | 458 | 7 | 1.7 | 490 | 5 | Q25175 | Q25175 haemochromatosis |
| 286 | 7 | 1.7 | 331 | 11 | Q91NM9 | Q91NM9 unclutered | 459 | 7 | 1.7 | 491 | 5 | Q25178 | Q25178 haemochromatosis |
| 287 | 7 | 1.7 | 332 | 5 | Q91NM9 | Q91NM9 unclutered | 460 | 7 | 1.7 | 492 | 5 | Q25181 | Q25181 haemochromatosis |
| 288 | 7 | 1.7 | 333 | 11 | Q91NM9 | Q91NM9 unclutered | 461 | 7 | 1.7 | 493 | 5 | Q25184 | Q25184 haemochromatosis |
| 289 | 7 | 1.7 | 333 | 11 | Q91NM9 | Q91NM9 unclutered | 462 | 7 | 1.7 | 494 | 5 | Q25187 | Q25187 haemochromatosis |
| 290 | 7 | 1.7 | 333 | 11 | Q91NM9 | Q91NM9 unclutered | 463 | 7 | 1.7 | 495 | 5 | Q25190 | Q25190 haemochromatosis |
| 291 | 7 | 1.7 | 333 | 11 | Q91NM9 | Q91NM9 unclutered | 464 | 7 | 1.7 | 496 | 5 | Q25193 | Q25193 haemochromatosis |
| 292 | 7 | 1.7 | 334 | 11 | Q91NM9 | Q91NM9 unclutered | 465 | 7 | 1.7 | 497 | 5 | Q25196 | Q25196 haemochromatosis |
| 293 | 7 | 1.7 | 334 | 11 | Q91NM9 | Q91NM9 unclutered | 466 | 7 | 1.7 | 498 | 5 | Q25199 | Q25199 haemochromatosis |
| 294 | 7 | 1.7 | 337 | 5 | Q91NM9 | Q91NM9 unclutered | 467 | 7 | 1.7 | 499 | 5 | Q25202 | Q25202 haemochromatosis |
| 295 | 7 | 1.7 | 337 | 12 | Q91NM9 | Q91NM9 unclutered | 468 | 7 | 1.7 | 500 | 5 | Q25205 | Q25205 haemochromatosis |
| 296 | 7 | 1.7 | 338 | 1 | Q91NM9 | Q91NM9 unclutered | 469 | 7 | 1.7 | 501 | 5 | Q25208 | Q25208 haemochromatosis |
| 297 | 7 | 1.7 | 338 | 10 | Q91NM9 | Q91NM9 unclutered | 470 | 7 | 1.7 | 502 | 5 | Q25211 | Q25211 haemochromatosis |
| 298 | 7 | 1.7 | 338 | 10 | Q91NM9 | Q91NM9 unclutered | 471 | 7 | 1.7 | 503 | 5 | Q25214 | Q25214 haemochromatosis |
| 299 | 7 | 1.7 | 339 | 4 | Q91NM9 | Q91NM9 unclutered | 472 | 7 | 1.7 | 504 | 5 | Q25217 | Q25217 haemochromatosis |
| 300 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 473 | 7 | 1.7 | 505 | 5 | Q25220 | Q25220 haemochromatosis |
| 301 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 474 | 7 | 1.7 | 506 | 5 | Q25223 | Q25223 haemochromatosis |
| 302 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 475 | 7 | 1.7 | 507 | 5 | Q25226 | Q25226 haemochromatosis |
| 303 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 476 | 7 | 1.7 | 508 | 5 | Q25229 | Q25229 haemochromatosis |
| 304 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 477 | 7 | 1.7 | 509 | 5 | Q25232 | Q25232 haemochromatosis |
| 305 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 478 | 7 | 1.7 | 510 | 5 | Q25235 | Q25235 haemochromatosis |
| 306 | 7 | 1.7 | 340 | 5 | Q91NM9 | Q91NM9 unclutered | 479 | 7 | 1.7 | 511 | 5 | Q25238 | Q25238 haemochromatosis |
| 307 | 7 | 1.7 | 340 | 11 | Q91NM9 | Q91NM9 unclutered | 480 | 7 | 1.7 | 512 | 5 | Q25241 | Q25241 haemochromatosis |
| 308 | 7 | 1.7 | 341 | 5 | Q91NM9 | Q91NM9 unclutered | 481 | 7 | 1.7 | 513 | 5 | Q25244 | Q25244 haemochromatosis |

| | | | | | | |
|-----|---|-----|-----|----|--------|----------------------|
| 528 | 6 | 1.5 | 132 | 17 | Q9PEK0 | Q9PEK0 acetophenone |
| 529 | 6 | 1.5 | 133 | 10 | Q42212 | Q42212 arabidopsis |
| 530 | 6 | 1.5 | 133 | 11 | Q63232 | Q63232 ratfus novy |
| 531 | 6 | 1.5 | 135 | 16 | Q9AMP2 | Q9AMP2 ribitolum |
| 532 | 6 | 1.5 | 136 | 10 | Q9ZNV5 | Q9ZNV5 saccharo core |
| 533 | 6 | 1.5 | 136 | 17 | Q28432 | Q28432 arabidopsis |
| 534 | 6 | 1.5 | 136 | 17 | Q9H107 | Q9H107 thermophilum |
| 535 | 6 | 1.5 | 137 | 5 | Q93952 | Q93952 arabidopsis |
| 536 | 6 | 1.5 | 137 | 5 | Q9B117 | Q9B117 paragonimus |
| 537 | 6 | 1.5 | 137 | 10 | Q46324 | Q46324 arabidopsis |
| 538 | 6 | 1.5 | 137 | 16 | Q9C8X6 | Q9C8X6 lactococcus |
| 539 | 6 | 1.5 | 137 | 16 | Q97010 | Q97010 arabidopsis |
| 540 | 6 | 1.5 | 138 | 8 | Q9S886 | Q9S886 guillardia |
| 541 | 6 | 1.5 | 138 | 8 | Q9FAB0 | Q9FAB0 vibrio hary |
| 542 | 6 | 1.5 | 139 | 2 | Q9FAP2 | Q9FAP2 vibrio hary |
| 543 | 6 | 1.5 | 139 | 2 | Q9FAP9 | Q9FAP9 vibrio hary |
| 544 | 6 | 1.5 | 139 | 2 | Q9FAB8 | Q9FAB8 vibrio hary |
| 545 | 6 | 1.5 | 139 | 2 | Q9FAB7 | Q9FAB7 vibrio hary |
| 546 | 6 | 1.5 | 139 | 2 | Q9FAB4 | Q9FAB4 vibrio hary |
| 547 | 6 | 1.5 | 139 | 2 | Q9FAB2 | Q9FAB2 vibrio hary |
| 548 | 6 | 1.5 | 139 | 2 | Q9FAB2 | Q9FAB2 vibrio hary |
| 549 | 6 | 1.5 | 139 | 2 | Q9FAP9 | Q9FAP9 vibrio hary |
| 550 | 6 | 1.5 | 139 | 2 | Q9FAP7 | Q9FAP7 vibrio hary |
| 551 | 6 | 1.5 | 139 | 2 | Q9FAP7 | Q9FAP7 vibrio hary |
| 552 | 6 | 1.5 | 139 | 2 | Q9FAP7 | Q9FAP7 vibrio hary |
| 553 | 6 | 1.5 | 139 | 2 | Q9FAP7 | Q9FAP7 vibrio hary |
| 554 | 6 | 1.5 | 139 | 10 | Q9FAP3 | Q9FAP3 arabidopsis |
| 555 | 6 | 1.5 | 140 | 5 | Q9FQ12 | Q9FQ12 drosophila |
| 556 | 6 | 1.5 | 140 | 5 | Q9FQ12 | Q9FQ12 drosophila |
| 557 | 6 | 1.5 | 141 | 15 | Q87056 | Q87056 simian foam |
| 558 | 6 | 1.5 | 142 | 4 | Q9F114 | Q9F114 citrobacter |
| 559 | 6 | 1.5 | 142 | 4 | Q9F114 | Q9F114 citrobacter |
| 560 | 6 | 1.5 | 142 | 16 | Q9C8X0 | Q9C8X0 homo sapiens |
| 561 | 6 | 1.5 | 143 | 12 | Q9F154 | Q9F154 arabidopsis |
| 562 | 6 | 1.5 | 143 | 12 | Q9F154 | Q9F154 arabidopsis |
| 563 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 564 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 565 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 566 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 567 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 568 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 569 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 570 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 571 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 572 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 573 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 574 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 575 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 576 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 577 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 578 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 579 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 580 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 581 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 582 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 583 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 584 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 585 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 586 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 587 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 588 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 589 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 590 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 591 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 592 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 593 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 594 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 595 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 596 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 597 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 598 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 599 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |
| 600 | 6 | 1.5 | 144 | 16 | Q9F154 | Q9F154 arabidopsis |

| | | | | | | | | | | | | | |
|-----|---|-----|-----|----|--------|---------------------|-----|---|-----|-----|----|--------|--------------------|
| 830 | 6 | 1.5 | 228 | 2 | Q9XYV4 | Q9XYV4 chitinella c | 894 | 6 | 1.5 | 255 | 14 | Q91290 | Q91290 pseudomonas |
| 821 | 6 | 1.5 | 228 | 2 | Q9A1M2 | Q9A1M2 pasteurella | 894 | 6 | 1.5 | 255 | 14 | Q91290 | Q91290 pseudomonas |
| 822 | 6 | 1.5 | 228 | 4 | Q9U9S5 | Q9U9S5 homo sapien | 895 | 6 | 1.5 | 256 | 17 | Q9H1B2 | Q9H1B2 thermoblast |
| 823 | 6 | 1.5 | 228 | 15 | Q9EYJ3 | Q9EYJ3 deinoxococ | 895 | 6 | 1.5 | 256 | 17 | Q9H1B2 | Q9H1B2 thermoblast |
| 824 | 6 | 1.5 | 229 | 17 | Q9Q9T4 | Q9Q9T4 salitribus | 897 | 6 | 1.5 | 257 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 825 | 6 | 1.5 | 229 | 9 | Q9MC82 | Q9MC82 bacteriophag | 898 | 6 | 1.5 | 257 | 10 | Q91597 | Q91597 putidaurum |
| 826 | 6 | 1.5 | 229 | 10 | Q9Q9R7 | Q9Q9R7 sea rays (m | 899 | 6 | 1.5 | 257 | 12 | Q91493 | Q91493 putidaurum |
| 827 | 6 | 1.5 | 230 | 16 | Q9Q9K3 | Q9Q9K3 chloridium | 900 | 6 | 1.5 | 257 | 16 | Q91493 | Q91493 putidaurum |
| 828 | 6 | 1.5 | 230 | 16 | Q9Q9Y2 | Q9Q9Y2 rhizobium m | 901 | 6 | 1.5 | 257 | 16 | Q91493 | Q91493 putidaurum |
| 829 | 6 | 1.5 | 231 | 4 | Q9A517 | Q9A517 homo sapien | 902 | 6 | 1.5 | 257 | 16 | Q91493 | Q91493 putidaurum |
| 830 | 6 | 1.5 | 231 | 9 | Q9A5M1 | Q9A5M1 lactococcus | 903 | 6 | 1.5 | 258 | 2 | Q91290 | Q91290 pseudomonas |
| 831 | 6 | 1.5 | 231 | 10 | Q9A5V7 | Q9A5V7 arabidopsis | 904 | 6 | 1.5 | 258 | 2 | Q91290 | Q91290 pseudomonas |
| 832 | 6 | 1.5 | 231 | 13 | Q91168 | Q91168 xenopus lae | 905 | 6 | 1.5 | 258 | 2 | Q91290 | Q91290 pseudomonas |
| 833 | 6 | 1.5 | 231 | 16 | Q9E632 | Q9E632 bacteriophag | 906 | 6 | 1.5 | 258 | 2 | Q91290 | Q91290 pseudomonas |
| 834 | 6 | 1.5 | 231 | 16 | Q9A104 | Q9A104 leptomema p | 907 | 6 | 1.5 | 258 | 12 | Q91493 | Q91493 putidaurum |
| 835 | 6 | 1.5 | 232 | 10 | Q9A1A1 | Q9A1A1 pyrococcus | 909 | 6 | 1.5 | 260 | 2 | Q9A1J3 | Q9A1J3 thermoblast |
| 836 | 6 | 1.5 | 232 | 17 | Q9H773 | Q9H773 thermoblast | 910 | 6 | 1.5 | 260 | 10 | Q9H1B2 | Q9H1B2 thermoblast |
| 837 | 6 | 1.5 | 232 | 17 | Q9H773 | Q9H773 thermoblast | 910 | 6 | 1.5 | 260 | 10 | Q9H1B2 | Q9H1B2 thermoblast |
| 838 | 6 | 1.5 | 232 | 17 | Q9H773 | Q9H773 thermoblast | 911 | 6 | 1.5 | 260 | 16 | Q9H1B2 | Q9H1B2 thermoblast |
| 839 | 6 | 1.5 | 233 | 12 | Q9EYJ3 | Q9EYJ3 deinoxococ | 912 | 6 | 1.5 | 260 | 16 | Q9H1B2 | Q9H1B2 thermoblast |
| 840 | 6 | 1.5 | 233 | 12 | Q9EYJ3 | Q9EYJ3 deinoxococ | 913 | 6 | 1.5 | 261 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 841 | 6 | 1.5 | 233 | 12 | Q91862 | Q91862 porfiroc cit | 914 | 6 | 1.5 | 261 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 842 | 6 | 1.5 | 233 | 12 | Q9A5M1 | Q9A5M1 lactococcus | 915 | 6 | 1.5 | 261 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 843 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 916 | 6 | 1.5 | 262 | 16 | Q9H1B2 | Q9H1B2 thermoblast |
| 844 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 917 | 6 | 1.5 | 262 | 16 | Q9H1B2 | Q9H1B2 thermoblast |
| 845 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 918 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 846 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 919 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 847 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 920 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 848 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 921 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 849 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 922 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 850 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 923 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 851 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 924 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 852 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 925 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 853 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 926 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 854 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 927 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 855 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 928 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 856 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 929 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 857 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 930 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 858 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 931 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 859 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 932 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 860 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 933 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 861 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 934 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 862 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 935 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 863 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 936 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 864 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 937 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 865 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 938 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 866 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 939 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 867 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 940 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 868 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 941 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 869 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 942 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 870 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 943 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 871 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 944 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 872 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 945 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 873 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 946 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 874 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 947 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 875 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 948 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 876 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 949 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 877 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 950 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 878 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 951 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 879 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 952 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 880 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 953 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 881 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 954 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 882 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 955 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 883 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 956 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 884 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 957 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 885 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 958 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 886 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 959 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 887 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 960 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 888 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 961 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 889 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 962 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 890 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 963 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 891 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 964 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |
| 892 | 6 | 1.5 | 233 | 12 | Q914M3 | Q914M3 porfiroc cit | 965 | 6 | 1.5 | 263 | 17 | Q9EYJ3 | Q9EYJ3 deinoxococ |

[illegible][illegible][illegible][illegible]

RN [1]
 EP SEQUENCE FROM N.A.
 RC STRAIN=16;
 RA Singh A., Sijwali P.S., Shenai R.P., Rosenthal P.J.;
 RT "Comparison of falcipain 2 gene sequence among multiple strains of
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Sijwali P.S., Singh A., Shenai R.P., Rosenthal P.J.;
 RT "Molecular Characterization of Falcipain 2: A Plasmodium falciparum
 RT Hemoglobinase."
 RI Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases
 DR EMBL: AF282977; AAF97807.1;
 DR FMRJ: AF284801; AAF97807.1;
 DR EMBL: AF282975; AAF97805.1;
 DR FMRJ: AF284766; AAF97806.1;
 DR HSSP: 060911; 1PH0.
 DR MEROPS: C01.046;
 EP InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprotease_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPA1N
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1
 KM Hydrolyase: Thiol protease.
 SQ SEQUENCE: 484 AA; 55992 MW; 5720603.944464 CRC64.

Query Match 2.78; Score 11; DR 5; Length 484;
 Best Local Similarity 100.0%; Prod. No. 0.017;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GRCGSCWAFS 204
 |||||
 DB 279 GRCGSCWAFS 289

RESULT 10
 Q9NAN2
 ID Q9NAN2 PRELIMINARY; PRT; 484 AA.
 AC Q9NAN2;
 DT 01-OCT-2000 (TREMBLrel.15, created)
 DT 01-OCT-2000 (TREMBLrel.15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel.19, last annotation update)
 DE FALCIPAIN 2 (FALCIPAIN-2).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Singh A., Sijwali P.S., Shenai R.P., Rosenthal P.J.;
 RT "Comparison of falcipain 2 gene sequence among multiple strains of
 RT Plasmodium falciparum."
 RI Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases
 DR EMBL: AF282977; AAF97807.1;
 DR FMRJ: AF284801; AAF97807.1;
 DR EMBL: AF282975; AAF97805.1;
 DR FMRJ: AF284766; AAF97806.1;
 DR HSSP: 060911; 1PH0.
 DR MEROPS: C01.046;
 EP InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprotease_site.
 DR Pfam: PF00112; Peptidase_C1; 1
 DR PRINTS: PR00705; PAPA1N
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1
 KM Hydrolyase: Thiol protease.
 SQ SEQUENCE: 484 AA; 55992 MW; 5720603.944464 CRC64.

Query Match 2.78; Score 11; DR 5; Length 484;
 Best Local Similarity 100.0%; Prod. No. 0.017;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GRCGSCWAFS 204
 |||||
 DB 279 GRCGSCWAFS 289

RESULT 11
 Q9NBD4
 ID Q9NBD4 PRELIMINARY; PRT; 484 AA.
 AC Q9NBD4;
 DT 01-OCT-2000 (TREMBLrel.15, created)
 DT 01-OCT-2000 (TREMBLrel.15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel.19, last annotation update)
 DE CYSTEINE PROTEASE FALCIPAIN-2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hanspal M.;
 RT "RNA cloning of a novel cysteine protease of Plasmodium falciparum."
 RI Biochem. Biophys. Acta 1493:242-245(2000).
 DR EMBL: AF251193; AAF65468.1;
 DR HSSP: 060911; 1PH0.
 DR MEROPS: C01.046;
 EP InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprotease_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPA1N
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1
 KM Hydrolyase: Protease; Thiol protease.
 SQ SEQUENCE: 484 AA; 55992 MW; E0034944CE1CB19 CRC64.

Query Match 2.78; Score 11; DR 5; Length 484;

Best Local Similarity 100.0%; Prod. No. 0.017;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GRCGSCWAFS 204
 |||||
 DB 279 GRCGSCWAFS 289

RESULT 12
 Q9NAN3
 ID Q9NAN3 PRELIMINARY; PRT; 484 AA.
 AC Q9NAN3;
 DT 01-OCT-2000 (TREMBLrel.15, created)
 DT 01-OCT-2000 (TREMBLrel.15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel.19, last annotation update)
 DE FALCIPAIN 2.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=11T5;
 RX MEDLINE=21091901; PubMed=11181388;
 RA Singh A., Rosenthal P.J.;
 RT "Comparison of Efficacies of Cysteine Protease Inhibitors against five
 RT strains of Plasmodium falciparum."
 RI Antimicrob. Agents Chemother. 46:446-451(2002).
 DR EMBL: AF282978; AAF97808.1;
 DR HSSP: 060911; 1PH0.
 DR MEROPS: C01.046;
 EP InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprotease_site.
 DR Pfam: PF00112; Peptidase_C1; 1.

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAV 206
|||||
DB 132 CGSCWAFSAV 141

RESULT 16

Q99KCO PRELIMINARY: PRT: 320 AA.
AC Q99KCO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE Cysteine protease 3.
OS Clonorchis sinensis.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Opisthorchiidae; Opisthorchiata; Opisthorchioidea; Opisthorchiidae;
OC Clonorchis.
NCBI_TaxID=79923;
RN [1]
RP SEQUENCE FROM N.A.
RA Yun P.-H., Chung T.-Y., Kong Y.-C., S. Y.
RT "Clonorchis sinensis protease 3 mRNA."
KL Submitted (MAY 2000) to the EMBL GenBank/GenBank databases.
DR EMBL: AF271091; AAK28439.1; -
DR HSSP: P25779; 1A1M.
DR MEROPS: C01.076; -
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000169; Thiolprol_act_site.
DR Pfam: PF00112; Peptidase_C1; 2.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN_1;
DR PROSITE: PS00139; THIOL_PROTEASE_CYS_1;
DR PROSITE: PS00639; THIOL_PROTEASE_HIS_1;
KM Hydrolyase; Protease; Thiol protease.
FT CHAIN 114..320
SQ SEQUENCE 320 AA: 36170 MW: 36170 MW: RDEDE7FA14551D06 CRG64;

Query Match: 2.5%; Score 10; DB 5; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAV 206
|||||
DB 135 CGSCWAFSAV 144

RESULT 17

Q95VA7 PRELIMINARY: PRT: 326 AA.
AC Q95VA7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE CAHEPSIN L.
GN CAT-LIG.
OS Fasciola gigantica.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Fasciolidae; Fasciolidae; Fasciolidae; Fasciolidae;
OC Fasciolidae; Fasciolidae; Fasciolidae; Fasciolidae;
NCBI_TaxID=46845;
RN [1]
RP SEQUENCE FROM N.A.
RA Schöen P., Meenen K., Grams K., Grams S.V., Korte G., Holmann A.;
RT "Molecular cloning of cathepsin L encoding genes from Fasciola
gigantica."
KL Submitted (Sep-2001) to the EMBL GenBank/GenBank databases.
DR EMBL: AF419329; AAL23917.1; -
SQ SEQUENCE 326 AA: 37457 MW: 70644 MW: 70644 MW: 70644 MW: 70644;

Query Match: 2.5%; Score 10; DB 5; Length 326;
Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAV 206
|||||
DB 129 CGSCWAFSAV 156

RESULT 18

Q99GN6 PRELIMINARY: PRT: 481 AA.
AC Q99GN6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE CAHEPSIN L LIKE CYSTEINE PROTEASE.
OS Leishmania chagasi.
OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmanidae;
OC Kinetoplastida; Kinetoplastida; Kinetoplastida; Kinetoplastida;
NCBI_TaxID=5669;
RN [1]
RP SEQUENCE FROM N.A.
RA Mundodi V., Samana A., Gedamu L.;
RT "Organization and functional analysis of a cysteine protease in
Leishmania chagasi."
KL Submitted (Sep-2001) to the EMBL GenBank/GenBank databases.
DR EMBL: AF212087; AAK16727.1; -
DR HSSP: P25779; 1A1M.
DR MEROPS: C01.076; -
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000169; Thiolprol_act_site.
DR Pfam: PF00112; Peptidase_C1; 2.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS_1;
KM Hydrolyase; Protease; Thiol protease.
SQ SEQUENCE 481 AA: 41362 MW: 41362 MW: 41362 MW: 41362 MW: 41362;

Query Match: 2.5%; Score 10; DB 5; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAV 206
|||||
DB 147 CGSCWAFSAV 156

RESULT 19

Q95WR6 PRELIMINARY: PRT: 394 AA.
AC Q95WR6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE CYSTEINE PROTEASE.
OS Leishmania donovani.
OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmanidae;
OC Kinetoplastida; Kinetoplastida; Kinetoplastida; Kinetoplastida;
NCBI_TaxID=5661;
RN [1]
RP SEQUENCE FROM N.A.
RA Mundodi V., Samana A., Gedamu L.;
RT "Isolation and characterization of cysteine protease genes from
Leishmania donovani."
KL Submitted (Sep-2001) to the EMBL GenBank/GenBank databases.
DR EMBL: AF309627; AAL09444.1; -
SQ SEQUENCE 394 AA: 42625 MW: 28942 MW: 28942 MW: 28942 MW: 28942;

Query Match: 2.5%; Score 10; DB 5; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSCWAFSAV 206
|||||
DB 147 CGSCWAFSAV 156

Query Match 2.28; Score 9; DB 5; Length 219;
Best Local Similarity 100.0%; Prod. No. 99;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 GSCSWAFSA 204
DB 21 GSCSWAFSA 29

RESULT 31

Q9XZM6

ID Q9XZM6 PRELIMINARY: PRT: 229 AA.

AC Q9XZM6: 01-NOV-1999 (TREMUREL_12, created)
DT 01-NOV-1999 (TREMUREL_12, last sequence update)
DI 01-DEC-2001 (TREMUREL_19, last annotation update)
DE CYSTEINE PROTEINASE A (FRAGMENT).
CN CPA.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: PAN (M96133.FP)
RX MEDLINE=2114442; PubMed=1134442;
RA Radaei S., Salimian A., Hashemi K., Schaff C., Boli S., Pasol M.;
RT "Identification of Leishmania major cysteine proteinases as targets of
the immune response in humans."
RI Mol. Biochem. Parasitol. 113:35-43(2001).
DR HSRP: A1130942; CAH3538.1;
DR HSRP: P25779; IAIM.
DR MEROPS: C01.076; .
DR InterPro: IPR000664; Peptidase_C1
DR InterPro: IPR000160; Thiolprot_act_site
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00130; THIOL_PROTEASE_CYS; 1
DR PROSITE: PS00640; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Thiol protease.
FT NON_TER 1
SQ SEQUENCE 229 AA. 2436 MW. 107741RTTAMT CE253.

Query Match

2.28; Score 9; DB 5; Length 229;
Best Local Similarity 100.0%; Prod. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GSCSWAFSA 205
DB 25 GSCSWAFSA 33

RESULT 32

Q26986

ID Q26986 PRELIMINARY: PRT: 270 AA.

AC Q26986: 01-NOV-1996 (TREMUREL_01, created)
DT 01-NOV-1996 (TREMUREL_01, last sequence update)
DI 01-DEC-2001 (TREMUREL_19, last annotation update)
DE TrpP2 PROTEIN (FRAGMENT).
CN TPCP2.
OS Trichomonas foetus (Trichomonas foetus).
OC Eukaryota; Parabasalida; Trichomonadida; Trichomonas.
OX NCBI_TaxID=5724;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: DL.
RX MEDLINE=96231873; PubMed=8627479;
RA Thorford J.W., Talbot J.A., Ikeda J.S., Corbett L.B.;
RT "Characterization of extracellular proteinases of Trichomonas
foetus."
RI J. Parasitol. 82:112-117(1996).
DR EMBL: U13154; AAB0163.1;

DR HSRP: P07584; IHOC.
DR MEROPS: C01.06A; .
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000160; Thiolprot_act_site.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00130; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00640; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Thiol protease.
FT NON_TER 1
SQ SEQUENCE 270 AA. 40394 MW. 41264RTTCAATP CE964.

Query Match

2.28; Score 9; DB 5; Length 270;
Best Local Similarity 100.0%; Prod. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GSCSWAFSA 205
DB 71 GSCSWAFSA 79

RESULT 33

Q27109

ID Q27109 PRELIMINARY: PRT: 292 AA.

AC Q27109: 01-NOV-1996 (TREMUREL_01, created)
DT 01-NOV-1996 (TREMUREL_01, last sequence update)
DI 01-DEC-2001 (TREMUREL_19, last annotation update)
DE CYSTEINE PROTEINASE, PUTATIVE (FRAGMENT).
CN CPA.
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalida; Trichomonadida; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: G3.
RX MEDLINE=95094621; PubMed=8605442;
RA Mathiasse D., Le Boudier M., Nachez A.M., Nachez M.;
RT "Identification and molecular cloning of four cysteine proteinase
genes from the pathogenic protozoan Trichomonas vaginalis."
RI Microbiology 140:2725-2735(1994).
DR HSRP: A77221; CA55118.1;
DR HSRP: 060911; IPR0.
DR MEROPS: C01.082; .
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000160; Thiolprot_act_site.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00640; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00130; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Thiol protease.
FT NON_TER 1
SQ SEQUENCE 292 AA. 42290 MW. 97216RTG9M06445 CE964.

Query Match

2.28; Score 9; DB 5; Length 292;
Best Local Similarity 100.0%; Prod. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GSCSWAFSA 205
DB 96 GSCSWAFSA 104

RESULT 34

Q96710

ID Q96710 PRELIMINARY: PRT: 310 AA.

AC Q96710: 01-DEC-2001 (TREMUREL_19, created)
DT 01-DEC-2001 (TREMUREL_19, last sequence update)
DI 01-DEC-2001 (TREMUREL_19, last annotation update)
DE CAHHEIN L. like (FRAGMENT).
CN CAHHEIN L. like (FRAGMENT).
OS Fasciola hepatica (liver fluke).
DR

FI CHAIN 99 314 Cysteine, PRE-THIAS, PUTATIVE.
 SO SEQUENCE 414 AA: 34641 MW: 477250441740 GPCB4:

Query Match 2.2% Score 9, DB 5, Length 314,
 Best Local Similarity 100.0%: Pred. No. 1.4:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 197 GCGSWAFSA 205
 DB 120 GCGSWAFSA 128

RESULT 38

ID Q26985 PRELIMINARY: PRT: 315 AA.

AC Q26985:
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE TROPIC PROTEIN.
 GN TROPIC.
 OS Trichomonas foetus (Trichomonas foetus).
 OC Eukaryota: Parabasalida: Trichomonadida: Trichomonas.
 OX NCBI_TaxID=5724:
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=66231873; PubMed=8627479;
 RA Thompson J W, Talbot J A, Iwaga J S, Corbett L R,
 RT "Characterization of extracellular proteinases of Trichomonas
 foetus".
 RL J. Parasitol. 82:112-117(1996).
 DR EMBL: U13153; AAB04162.1;
 DR HSSP: P07711; ICUL.
 DR MEROPS: C01.082; -.
 DR InterPro: IPR000666; Peptidase_C1
 DR InterPro: IPR000169; Thiolprol_act_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR PROSITE: PS00640; Thiol_protease_ASN; 1
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 KW Hydrolase; Thiol protease.
 FT CHAIN 100 315 POTENTIAL.
 ST SEQUENCE 315 AA: 35154 MW: 49489844444 GPCB4:

Query Match 2.2% Score 9, DB 5, Length 314,
 Best Local Similarity 100.0%: Pred. No. 1.4:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 197 GCGSWAFSA 205
 DB 121 GCGSWAFSA 129

RESULT 39

ID Q26564 PRELIMINARY: PRT: 317 AA.

AC Q26564:
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE CATHEPSIN L (FRAGMENT).
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota: Metazoa: Platyhelminthes: Trematoda: Digenea: Strigeidae;
 OC Schistosomatidae: Schistosomatidae: Schistosoma.
 OX NCBI_TaxID=6183:
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=96123374; PubMed=8577349;
 RA Michel A, Chongem H, Rosta M, Krickert M U, Kunz W,
 RT "Sequence, characterization and localization of a cysteine proteinase
 cathepsin L in Schistosoma mansoni".

FI MOI: Biochem. Parasitol. 7:17-18(1990).

DR EMBL: Z42529; CAAB3548.1;
 DR HSSP: P07711; ICUL.
 DR MEROPS: C01.044; -.
 DR InterPro: IPR000666; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprol_act_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 KW Hydrolase; Thiol protease.
 FT NON_TER 1
 FT CHAIN 102 317 CATHEPSIN L.
 ST SEQUENCE 317 AA: 36571 MW: 44630128072485 GPCB4:

Query Match 2.2% Score 9, DB 5, Length 317,
 Best Local Similarity 100.0%: Pred. No. 1.4:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 197 GCGSWAFSA 205
 DB 124 GCGSWAFSA 132

RESULT 40

ID Q9G062 PRELIMINARY: PRT: 422 AA.

AC Q9G062:
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE CATHEPSIN L (LIFE CYCLE) (TEMBLrel. 16, Last annotation update)
 GN CALL.
 OS Diabrotica virgifera virgifera (western corn rootworm).
 OC Eukaryota: Metazoa: Arthropoda: Insecta: Hexapoda: Insecta:
 OC Curculionida: Brachyura: Endopterygota: Coleoptera: Polyphaga:
 OC Curculioninae: Phytophaga: Chrysomeloidae: Chrysomelidae: Galerucinae:
 OC Diabrotinae.
 OX NCBI_TaxID=50300:
 RN [1]
 RP SEQUENCE FROM N.A.

RA Brown H, Shode R E, Zhan Salzman K, Pithers M F, Muthukrishnan
 RA Brown R A, Hasegawa P M;
 RT "A plant defensive cystatin (scorpionatin) targets cathepsin L-like
 digestive cysteine proteinases (PCPases) in the larval midgut of
 western corn rootworm (Diabrotica virgifera virgifera)".
 RL J. Biol. Chem. 273:11102-11107 (1998).
 DR EMBL: AF190653; AAC17127.1;
 DR HSSP: 060911; IFHO.

DR MEROPS: C01.08A; -.
 DR InterPro: IPR000666; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprol_act_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPAIN.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 KW Hydrolase; Thiol protease.
 FT NON_TER 1
 ST SEQUENCE 422 AA: 46673 MW: 85040144404157 GPCB4:

Query Match 2.2% Score 9, DB 5, Length 422,
 Best Local Similarity 100.0%: Pred. No. 1.4:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 197 GCGSWAFSA 205
 DB 129 GCGSWAFSA 137

RESULT 41

ID Q96705 PRELIMINARY: PRT: 422 AA.

AC Q96705:

QY 147 CGSCWAFSA 205
|||||||
DB 135 CGSCWAFSA 143

Search completed: October 9, 2002, 11:48:30
Job time : 55 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 11:16:55 : Search time 19 seconds

(without alignments)
515,509 Million cell updates/sec

Title: US-09-598-062-4

Perfect score: 401

Sequence: 1 MOTGNVPPHVEYISAPYIA.....KKGTGGLVYVYVHNSI 401

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 231628 seqs, 2412594 residues

Word size: 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2099090990

Post-processing: Listing first 1000 summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | length DB | ID | Description |
|------------|-------|---------------|-----------|-------------------|--------------------|
| 1 | 401 | 100.0 | 401 | US-08-827-171P-4 | Sequence 4, Aff11 |
| 2 | 326 | 81.4 | 401 | US-08-827-171P-4 | Sequence 6, Aff11 |
| 3 | 176 | 43.6 | 401 | US-08-827-171P-4 | Sequence 7, Aff11 |
| 4 | 12 | 3.0 | 344 | US-08-453-924-10 | Sequence 10, Aff11 |
| 5 | 10 | 2.5 | 212 | US-08-860-255A-4 | Sequence 1, Aff11 |
| 6 | 10 | 2.5 | 314 | US-08-981-957B-1* | Sequence 13, Aff11 |
| 7 | 10 | 2.5 | 345 | US-09-120-365-69 | Sequence 73, Aff11 |
| 8 | 10 | 2.5 | 345 | US-09-515-039-66 | Sequence 69, Aff11 |
| 9 | 10 | 2.5 | 345 | US-08-827-171P-4 | Sequence 6, Aff11 |
| 10 | 10 | 2.5 | 345 | US-09-120-365-69 | Sequence 73, Aff11 |
| 11 | 10 | 2.5 | 443 | US-09-120-365-69 | Sequence 69, Aff11 |
| 12 | 10 | 2.5 | 443 | US-09-515-039-66 | Sequence 69, Aff11 |
| 13 | 9 | 2.2 | 326 | US-09-515-039-66 | Sequence 67, Aff11 |
| 14 | 9 | 2.2 | 326 | US-09-515-039-66 | Sequence 67, Aff11 |
| 15 | 9 | 2.2 | 343 | US-08-209-097A-9 | Sequence 9, Aff11 |
| 16 | 9 | 2.2 | 343 | US-08-883-526-4 | Sequence 3, Aff11 |
| 17 | 9 | 2.2 | 343 | US-09-120-365-69 | Sequence 73, Aff11 |
| 18 | 9 | 2.2 | 343 | US-09-515-039-66 | Sequence 69, Aff11 |
| 19 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 20 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 21 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 22 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 23 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 24 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 25 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 26 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |
| 27 | 9 | 2.2 | 343 | US-08-860-255A-4 | Sequence 65, Aff11 |

| | | | | | |
|-----|---|-----|-----|-------------------|--------------------|
| 28 | 9 | 2.2 | 457 | US-09-120-365-69 | Sequence 72, Aff11 |
| 29 | 9 | 2.2 | 457 | US-09-515-039-66 | Sequence 72, Aff11 |
| 30 | 8 | 2.0 | 11 | US-09-119-361-15 | Sequence 1, Aff11 |
| 31 | 8 | 2.0 | 14 | US-08-436-308A-15 | Sequence 15, Aff11 |
| 32 | 8 | 2.0 | 14 | US-09-490-931-15 | Sequence 15, Aff11 |
| 33 | 8 | 2.0 | 27 | US-08-482-142-10 | Sequence 10, Aff11 |
| 34 | 8 | 2.0 | 27 | US-08-482-142-73 | Sequence 73, Aff11 |
| 35 | 8 | 2.0 | 27 | US-08-478-572-10 | Sequence 10, Aff11 |
| 36 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 37 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 38 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 39 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 40 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 41 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 42 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 43 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 44 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 45 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 46 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 47 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 48 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 49 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 50 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 51 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 52 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 53 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 54 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 55 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 56 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 57 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 58 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 59 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 60 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 61 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 62 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 63 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 64 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 65 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 66 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 67 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 68 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 69 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 70 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 71 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 72 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 73 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 74 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 75 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 76 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 77 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 78 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 79 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 80 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 81 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 82 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 83 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 84 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 85 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 86 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 87 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 88 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 89 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 90 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 91 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 92 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 93 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 94 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 95 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 96 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 97 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 98 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 99 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |
| 100 | 8 | 2.0 | 27 | US-08-478-572-73 | Sequence 73, Aff11 |

| | | | | | | | | | | | | | | | |
|-----|---|---|---|----|---|--------------------|--------------------|-----|---|---|---|----|---|-------------------|--------------------|
| 539 | 5 | 1 | 2 | 15 | 2 | US-09-056-849-19 | Sequence 18, App1 | 612 | 5 | 1 | 2 | 23 | 4 | US-08-468-740-14 | Sequence 14, App1 |
| 540 | 5 | 1 | 2 | 15 | 3 | US-08-948-547-18 | Sequence 18, App1 | 613 | 5 | 1 | 2 | 23 | 4 | US-08-484-296-40 | Sequence 40, App1 |
| 541 | 5 | 1 | 2 | 15 | 4 | US-08-602-999A-345 | Sequence 345, App | 614 | 5 | 1 | 2 | 23 | 4 | US-08-484-296-104 | Sequence 104, App |
| 542 | 5 | 1 | 2 | 15 | 4 | US-09-248-588-25 | Sequence 25, App1 | 615 | 5 | 1 | 2 | 23 | 4 | US-07-791-945E-10 | Sequence 2, App1 |
| 543 | 5 | 1 | 2 | 15 | 4 | US-08-160-604-115 | Sequence 115, App | 616 | 5 | 1 | 2 | 23 | 4 | US-08-141-121-22 | Sequence 22, App1 |
| 544 | 5 | 1 | 2 | 15 | 4 | US-09-406-781-4 | Sequence 4, App1 | 617 | 5 | 1 | 2 | 23 | 4 | US-09-541-992-22 | Sequence 22, App1 |
| 545 | 5 | 1 | 2 | 15 | 4 | US-08-956-653A-18 | Sequence 18, App1 | 618 | 5 | 1 | 2 | 23 | 4 | US-08-482-112-22 | Sequence 22, App1 |
| 546 | 5 | 1 | 2 | 15 | 4 | US-09-025-596-20 | Sequence 20, App1 | 619 | 5 | 1 | 2 | 23 | 4 | US-08-482-112-88 | Sequence 88, App1 |
| 547 | 5 | 1 | 2 | 15 | 4 | US-09-194-062-10 | Sequence 10, App1 | 620 | 5 | 1 | 2 | 23 | 4 | US-08-712-113A-20 | Sequence 20, App1 |
| 548 | 5 | 1 | 2 | 16 | 3 | US-08-335-198-4 | Sequence 4, App1 | 621 | 5 | 1 | 2 | 23 | 4 | US-08-478-572-26 | Sequence 26, App1 |
| 549 | 5 | 1 | 2 | 16 | 3 | US-08-782-180-45 | Sequence 45, App1 | 622 | 5 | 1 | 2 | 23 | 4 | US-08-478-572-89 | Sequence 89, App1 |
| 550 | 5 | 1 | 2 | 16 | 4 | US-08-160-504-3 | Sequence 3, App1 | 623 | 5 | 1 | 2 | 23 | 4 | US-07-711-451A-22 | Sequence 22, App1 |
| 551 | 5 | 1 | 2 | 16 | 4 | US-08-934-211-45 | Sequence 45, App1 | 624 | 5 | 1 | 2 | 23 | 4 | US-08-484-296-25 | Sequence 25, App1 |
| 552 | 5 | 1 | 2 | 16 | 4 | US-08-913-162-9 | Sequence 9, App1 | 625 | 5 | 1 | 2 | 23 | 4 | US-08-484-296-89 | Sequence 89, App1 |
| 553 | 5 | 1 | 2 | 17 | 1 | US-08-520-977A-10 | Sequence 10, App1 | 626 | 5 | 1 | 2 | 23 | 4 | US-08-484-296-104 | Sequence 104, App1 |
| 554 | 5 | 1 | 2 | 17 | 1 | US-07-791-945E-3 | Sequence 3, App1 | 627 | 5 | 1 | 2 | 23 | 4 | US-08-318-193-24 | Sequence 24, App1 |
| 555 | 5 | 1 | 2 | 18 | 1 | US-08-374-654C-89 | Sequence 89, App1 | 628 | 5 | 1 | 2 | 23 | 4 | US-08-248-839C-95 | Sequence 95, App1 |
| 556 | 5 | 1 | 2 | 18 | 1 | US-08-345-503-17 | Sequence 17, App1 | 629 | 5 | 1 | 2 | 23 | 4 | US-08-592-500-34 | Sequence 34, App1 |
| 557 | 5 | 1 | 2 | 19 | 1 | US-08-751-305-13 | Sequence 13, App1 | 630 | 5 | 1 | 2 | 23 | 4 | US-08-195-004-33 | Sequence 33, App1 |
| 558 | 5 | 1 | 2 | 19 | 2 | US-08-462-436-15 | Sequence 15, App1 | 631 | 5 | 1 | 2 | 23 | 4 | US-09-125-244-1 | Sequence 1, App1 |
| 559 | 5 | 1 | 2 | 19 | 3 | US-08-462-436-18 | Sequence 18, App1 | 632 | 5 | 1 | 2 | 23 | 4 | US-09-218-599-72 | Sequence 72, App1 |
| 560 | 5 | 1 | 2 | 19 | 3 | US-08-868-544-18 | Sequence 18, App1 | 633 | 5 | 1 | 2 | 23 | 4 | US-07-642-245-290 | Sequence 290, App1 |
| 561 | 5 | 1 | 2 | 19 | 4 | US-08-445-375-15 | Sequence 15, App1 | 634 | 5 | 1 | 2 | 23 | 4 | US-07-642-245-294 | Sequence 294, App1 |
| 562 | 5 | 1 | 2 | 19 | 4 | US-08-465-325-18 | Sequence 18, App1 | 635 | 5 | 1 | 2 | 23 | 4 | US-07-942-245-416 | Sequence 416, App1 |
| 563 | 5 | 1 | 2 | 19 | 4 | US-09-434-423-10 | Sequence 10, App1 | 636 | 5 | 1 | 2 | 23 | 4 | US-07-942-245-416 | Sequence 416, App1 |
| 564 | 5 | 1 | 2 | 20 | 4 | US-09-406-781-1 | Sequence 1, App1 | 637 | 5 | 1 | 2 | 23 | 4 | US-09-942-245-416 | Sequence 416, App1 |
| 565 | 5 | 1 | 2 | 20 | 4 | US-09-406-781-1 | Sequence 1, App1 | 638 | 5 | 1 | 2 | 23 | 4 | US-09-942-245-416 | Sequence 416, App1 |
| 566 | 5 | 1 | 2 | 20 | 6 | 5262298-2 | Sequence 2, App1 | 639 | 5 | 1 | 2 | 23 | 4 | US-09-942-245-416 | Sequence 416, App1 |
| 567 | 5 | 1 | 2 | 20 | 6 | 5262298-1 | Sequence 1, App1 | 640 | 5 | 1 | 2 | 23 | 4 | US-09-942-245-416 | Sequence 416, App1 |
| 568 | 5 | 1 | 2 | 21 | 1 | US-08-555-579-1 | Sequence 1, App1 | 641 | 5 | 1 | 2 | 23 | 4 | US-09-942-245-416 | Sequence 416, App1 |
| 569 | 5 | 1 | 2 | 21 | 3 | US-08-651-146C-36 | Sequence 36, App1 | 642 | 5 | 1 | 2 | 23 | 4 | US-08-651-146C-74 | Sequence 74, App1 |
| 570 | 5 | 1 | 2 | 21 | 3 | US-08-651-146C-41 | Sequence 41, App1 | 643 | 5 | 1 | 2 | 23 | 4 | US-08-651-146C-74 | Sequence 74, App1 |
| 571 | 5 | 1 | 2 | 22 | 2 | US-08-498-859A-6 | Sequence 6, App1 | 644 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 572 | 5 | 1 | 2 | 22 | 2 | US-08-742-452A-21 | Sequence 21, App1 | 645 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 573 | 5 | 1 | 2 | 22 | 3 | US-08-940-095-133 | Sequence 133, App1 | 646 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 574 | 5 | 1 | 2 | 22 | 3 | US-08-940-095-173 | Sequence 173, App1 | 647 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 575 | 5 | 1 | 2 | 22 | 3 | US-08-940-095-175 | Sequence 175, App1 | 648 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 576 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-187 | Sequence 187, App1 | 649 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 577 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 650 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 578 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 651 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 579 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 652 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 580 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 653 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 581 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 654 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 582 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 655 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 583 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 656 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 584 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 657 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 585 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 658 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 586 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 659 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 587 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 660 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 588 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 661 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 589 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 662 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 590 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 663 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 591 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 664 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 592 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 665 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 593 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 666 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 594 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 667 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 595 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 668 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 596 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 669 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 597 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 670 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 598 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 671 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 599 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 672 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 600 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 673 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 601 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 674 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 602 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 675 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 603 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 676 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 604 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 677 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 605 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 678 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 606 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 679 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 607 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 680 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 608 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 681 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 609 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 682 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 610 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 683 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |
| 611 | 5 | 1 | 2 | 22 | 4 | US-08-940-095-133 | Sequence 133, App1 | 684 | 5 | 1 | 2 | 23 | 4 | US-08-498-859A-6 | Sequence 6, App1 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|---|-----|----|---|-------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|
| 831 | 5 | 1.2 | 86 | 3 | US-08-645-8600-53 | Sequence 53, April | 534 | 5 | 1.2 | 106 | 3 | US-09-083-351-15 | Sequence 15, April | 541 | 5 | 1.2 | 107 | 4 | US-09-083-352-16 | Sequence 16, April | 542 | 5 | 1.2 | 108 | 5 | US-09-083-353-17 | Sequence 17, April | 543 | 5 | 1.2 | 109 | 6 | US-09-083-354-18 | Sequence 18, April | 544 | 5 | 1.2 | 110 | 7 | US-09-083-355-19 | Sequence 19, April | 545 | 5 | 1.2 | 111 | 8 | US-09-083-356-20 | Sequence 20, April | 546 | 5 | 1.2 | 112 | 9 | US-09-083-357-21 | Sequence 21, April | 547 | 5 | 1.2 | 113 | 10 | US-09-083-358-22 | Sequence 22, April | 548 | 5 | 1.2 | 114 | 11 | US-09-083-359-23 | Sequence 23, April | 549 | 5 | 1.2 | 115 | 12 | US-09-083-360-24 | Sequence 24, April | 550 | 5 | 1.2 | 116 | 13 | US-09-083-361-25 | Sequence 25, April | 551 | 5 | 1.2 | 117 | 14 | US-09-083-362-26 | Sequence 26, April | 552 | 5 | 1.2 | 118 | 15 | US-09-083-363-27 | Sequence 27, April | 553 | 5 | 1.2 | 119 | 16 | US-09-083-364-28 | Sequence 28, April | 554 | 5 | 1.2 | 120 | 17 | US-09-083-365-29 | Sequence 29, April | 555 | 5 | 1.2 | 121 | 18 | US-09-083-366-30 | Sequence 30, April | 556 | 5 | 1.2 | 122 | 19 | US-09-083-367-31 | Sequence 31, April | 557 | 5 | 1.2 | 123 | 20 | US-09-083-368-32 | Sequence 32, April | 558 | 5 | 1.2 | 124 | 21 | US-09-083-369-33 | Sequence 33, April | 559 | 5 | 1.2 | 125 | 22 | US-09-083-370-34 | Sequence 34, April | 560 | 5 | 1.2 | 126 | 23 | US-09-083-371-35 | Sequence 35, April | 561 | 5 | 1.2 | 127 | 24 | US-09-083-372-36 | Sequence 36, April | 562 | 5 | 1.2 | 128 | 25 | US-09-083-373-37 | Sequence 37, April | 563 | 5 | 1.2 | 129 | 26 | US-09-083-374-38 | Sequence 38, April | 564 | 5 | 1.2 | 130 | 27 | US-09-083-375-39 | Sequence 39, April | 565 | 5 | 1.2 | 131 | 28 | US-09-083-376-40 | Sequence 40, April | 566 | 5 | 1.2 | 132 | 29 | US-09-083-377-41 | Sequence 41, April | 567 | 5 | 1.2 | 133 | 30 | US-09-083-378-42 | Sequence 42, April | 568 | 5 | 1.2 | 134 | 31 | US-09-083-379-43 | Sequence 43, April | 569 | 5 | 1.2 | 135 | 32 | US-09-083-380-44 | Sequence 44, April | 570 | 5 | 1.2 | 136 | 33 | US-09-083-381-45 | Sequence 45, April | 571 | 5 | 1.2 | 137 | 34 | US-09-083-382-46 | Sequence 46, April | 572 | 5 | 1.2 | 138 | 35 | US-09-083-383-47 | Sequence 47, April | 573 | 5 | 1.2 | 139 | 36 | US-09-083-384-48 | Sequence 48, April | 574 | 5 | 1.2 | 140 | 37 | US-09-083-385-49 | Sequence 49, April | 575 | 5 | 1.2 | 141 | 38 | US-09-083-386-50 | Sequence 50, April | 576 | 5 | 1.2 | 142 | 39 | US-09-083-387-51 | Sequence 51, April | 577 | 5 | 1.2 | 143 | 40 | US-09-083-388-52 | Sequence 52, April | 578 | 5 | 1.2 | 144 | 41 | US-09-083-389-53 | Sequence 53, April | 579 | 5 | 1.2 | 145 | 42 | US-09-083-390-54 | Sequence 54, April | 580 | 5 | 1.2 | 146 | 43 | US-09-083-391-55 | Sequence 55, April | 581 | 5 | 1.2 | 147 | 44 | US-09-083-392-56 | Sequence 56, April | 582 | 5 | 1.2 | 148 | 45 | US-09-083-393-57 | Sequence 57, April | 583 | 5 | 1.2 | 149 | 46 | US-09-083-394-58 | Sequence 58, April | 584 | 5 | 1.2 | 150 | 47 | US-09-083-395-59 | Sequence 59, April | 585 | 5 | 1.2 | 151 | 48 | US-09-083-396-60 | Sequence 60, April | 586 | 5 | 1.2 | 152 | 49 | US-09-083-397-61 | Sequence 61, April | 587 | 5 | 1.2 | 153 | 50 | US-09-083-398-62 | Sequence 62, April | 588 | 5 | 1.2 | 154 | 51 | US-09-083-399-63 | Sequence 63, April | 589 | 5 | 1.2 | 155 | 52 | US-09-083-400-64 | Sequence 64, April | 590 | 5 | 1.2 | 156 | 53 | US-09-083-401-65 | Sequence 65, April | 591 | 5 | 1.2 | 157 | 54 | US-09-083-402-66 | Sequence 66, April | 592 | 5 | 1.2 | 158 | 55 | US-09-083-403-67 | Sequence 67, April | 593 | 5 | 1.2 | 159 | 56 | US-09-083-404-68 | Sequence 68, April | 594 | 5 | 1.2 | 160 | 57 | US-09-083-405-69 | Sequence 69, April | 595 | 5 | 1.2 | 161 | 58 | US-09-083-406-70 | Sequence 70, April | 596 | 5 | 1.2 | 162 | 59 | US-09-083-407-71 | Sequence 71, April | 597 | 5 | 1.2 | 163 | 60 | US-09-083-408-72 | Sequence 72, April | 598 | 5 | 1.2 | 164 | 61 | US-09-083-409-73 | Sequence 73, April | 599 | 5 | 1.2 | 165 | 62 | US-09-083-410-74 | Sequence 74, April | 600 | 5 | 1.2 | 166 | 63 | US-09-083-411-75 | Sequence 75, April | 601 | 5 | 1.2 | 167 | 64 | US-09-083-412-76 | Sequence 76, April | 602 | 5 | 1.2 | 168 | 65 | US-09-083-413-77 | Sequence 77, April | 603 | 5 | 1.2 | 169 | 66 | US-09-083-414-78 | Sequence 78, April | 604 | 5 | 1.2 | 170 | 67 | US-09-083-415-79 | Sequence 79, April | 605 | 5 | 1.2 | 171 | 68 | US-09-083-416-80 | Sequence 80, April | 606 | 5 | 1.2 | 172 | 69 | US-09-083-417-81 | Sequence 81, April | 607 | 5 | 1.2 | 173 | 70 | US-09-083-418-82 | Sequence 82, April | 608 | 5 | 1.2 | 174 | 71 | US-09-083-419-83 | Sequence 83, April | 609 | 5 | 1.2 | 175 | 72 | US-09-083-420-84 | Sequence 84, April | 610 | 5 | 1.2 | 176 | 73 | US-09-083-421-85 | Sequence 85, April | 611 | 5 | 1.2 | 177 | 74 | US-09-083-422-86 | Sequence 86, April | 612 | 5 | 1.2 | 178 | 75 | US-09-083-423-87 | Sequence 87, April | 613 | 5 | 1.2 | 179 | 76 | US-09-083-424-88 | Sequence 88, April | 614 | 5 | 1.2 | 180 | 77 | US-09-083-425-89 | Sequence 89, April | 615 | 5 | 1.2 | 181 | 78 | US-09-083-426-90 | Sequence 90, April | 616 | 5 | 1.2 | 182 | 79 | US-09-083-427-91 | Sequence 91, April | 617 | 5 | 1.2 | 183 | 80 | US-09-083-428-92 | Sequence 92, April | 618 | 5 | 1.2 | 184 | 81 | US-09-083-429-93 | Sequence 93, April | 619 | 5 | 1.2 | 185 | 82 | US-09-083-430-94 | Sequence 94, April | 620 | 5 | 1.2 | 186 | 83 | US-09-083-431-95 | Sequence 95, April | 621 | 5 | 1.2 | 187 | 84 | US-09-083-432-96 | Sequence 96, April | 622 | 5 | 1.2 | 188 | 85 | US-09-083-433-97 | Sequence 97, April | 623 | 5 | 1.2 | 189 | 86 | US-09-083-434-98 | Sequence 98, April | 624 | 5 | 1.2 | 190 | 87 | US-09-083-435-99 | Sequence 99, April | 625 | 5 | 1.2 | 191 | 88 | US-09-083-436-00 | Sequence 100, April | 626 | 5 | 1.2 | 192 | 89 | US-09-083-437-01 | Sequence 101, April | 627 | 5 | 1.2 | 193 | 90 | US-09-083-438-02 | Sequence 102, April | 628 | 5 | 1.2 | 194 | 91 | US-09-083-439-03 | Sequence 103, April | 629 | 5 | 1.2 | 195 | 92 | US-09-083-440-04 | Sequence 104, April | 630 | 5 | 1.2 | 196 | 93 | US-09-083-441-05 | Sequence 105, April | 631 | 5 | 1.2 | 197 | 94 | US-09-083-442-06 | Sequence 106, April | 632 | 5 | 1.2 | 198 | 95 | US-09-083-443-07 | Sequence 107, April | 633 | 5 | 1.2 | 199 | 96 | US-09-083-444-08 | Sequence 108, April | 634 | 5 | 1.2 | 200 | 97 | US-09-083-445-09 | Sequence 109, April | 635 | 5 | 1.2 | 201 | 98 | US-09-083-446-10 | Sequence 110, April | 636 | 5 | 1.2 | 202 | 99 | US-09-083-447-11 | Sequence 111, April | 637 | 5 | 1.2 | 203 | 00 | US-09-083-448-12 | Sequence 112, April | 638 | 5 | 1.2 | 204 | 01 | US-09-083-449-13 | Sequence 113, April | 639 | 5 | 1.2 | 205 | 02 | US-09-083-450-14 | Sequence 114, April | 640 | 5 | 1.2 | 206 | 03 | US-09-083-451-15 | Sequence 115, April | 641 | 5 | 1.2 | 207 | 04 | US-09-083-452-16 | Sequence 116, April | 642 | 5 | 1.2 | 208 | 05 | US-09-083-453-17 | Sequence 117, April | 643 | 5 | 1.2 | 209 | 06 | US-09-083-454-18 | Sequence 118, April | 644 | 5 | 1.2 | 210 | 07 | US-09-083-455-19 | Sequence 119, April | 645 | 5 | 1.2 | 211 | 08 | US-09-083-456-20 | Sequence 120, April | 646 | 5 | 1.2 | 212 | 09 | US-09-083-457-21 | Sequence 121, April | 647 | 5 | 1.2 | 213 | 10 | US-09-083-458-22 | Sequence 122, April | 648 | 5 | 1.2 | 214 | 11 | US-09-083-459-23 | Sequence 123, April | 649 | 5 | 1.2 | 215 | 12 | US-09-083-460-24 | Sequence 124, April | 650 | 5 | 1.2 | 216 | 13 | US-09-083-461-25 | Sequence 125, April | 651 | 5 | 1.2 | 217 | 14 | US-09-083-462-26 | Sequence 126, April | 652 | 5 | 1.2 | 218 | 15 | US-09-083-463-27 | Sequence 127, April | 653 | 5 | 1.2 | 219 | 16 | US-09-083-464-28 | Sequence 128, April | 654 | 5 | 1.2 | 220 | 17 | US-09-083-465-29 | Sequence 129, April | 655 | 5 | 1.2 | 221 | 18 | US-09-083-466-30 | Sequence 130, April | 656 | 5 | 1.2 | 222 | 19 | US-09-083-467-31 | Sequence 131, April | 657 | 5 | 1.2 | 223 | 20 | US-09-083-468-32 | Sequence 132, April | 658 | 5 | 1.2 | 224 | 21 | US-09-083-469-33 | Sequence 133, April | 659 | 5 | 1.2 | 225 | 22 | US-09-083-470-34 | Sequence 134, April | 660 | 5 | 1.2 | 226 | 23 | US-09-083-471-35 | Sequence 135, April | 661 | 5 | 1.2 | 227 | 24 | US-09-083-472-36 | Sequence 136, April | 662 | 5 | 1.2 | 228 | 25 | US-09-083-473-37 | Sequence 137, April | 663 | 5 | 1.2 | 229 | 26 | US-09-083-474-38 | Sequence 138, April | 664 | 5 | 1.2 | 230 | 27 | US-09-083-475-39 | Sequence 139, April | 665 | 5 | 1.2 | 231 | 28 | US-09-083-476-40 | Sequence 140, April | 666 | 5 | 1.2 | 232 | 29 | US-09-083-477-41 | Sequence 141, April | 667 | 5 | 1.2 | 233 | 30 | US-09-083-478-42 | Sequence 142, April | 668 | 5 | 1.2 | 234 | 31 | US-09-083-479-43 | Sequence 143, April | 669 | 5 | 1.2 | 235 | 32 | US-09-083-480-44 | Sequence 144, April | 670 | 5 | 1.2 | 236 | 33 | US-09-083-481-45 | Sequence 145, April | 671 | 5 | 1.2 | 237 | 34 | US-09-083-482-46 | Sequence 146, April | 672 | 5 | 1.2 | 238 | 35 | US-09-083-483-47 | Sequence 147, April | 673 | 5 | 1.2 | 239 | 36 | US-09-083-484-48 | Sequence 148, April | 674 | 5 | 1.2 | 240 | 37 | US-09-083-485-49 | Sequence 149, April | 675 | 5 | 1.2 | 241 | 38 | US-09-083-486-50 | Sequence 150, April | 676 | 5 | 1.2 | 242 | 39 | US-09-083-487-51 | Sequence 151, April | 677 | 5 | 1.2 | 243 | 40 | US-09-083-488-52 | Sequence 152, April | 678 | 5 | 1.2 | 244 | 41 | US-09-083-489-53 | Sequence 153, April | 679 | 5 | 1.2 | 245 | 42 | US-09-083-490-54 | Sequence 154, April | 680 | 5 | 1.2 | 246 | 43 | US-09-083-491-55 | Sequence 155, April | 681 | 5 | 1.2 | 247 | 44 | US-09-083-492-56 | Sequence 156, April | 682 | 5 | 1.2 | 248 | 45 | US-09-083-493-57 | Sequence 157, April | 683 | 5 | 1.2 | 249 | 46 | US-09-083-494-58 | Sequence 158, April | 684 | 5 | 1.2 | 250 | 47 | US-09-083-495-59 | Sequence 159, April | 685 | 5 | 1.2 | 251 | 48 | US-09-083-496-60 | Sequence 160, April | 686 | 5 | 1.2 | 252 | 49 | US-09-083-497-61 | Sequence 161, April | 687 | 5 | 1.2 | 253 | 50 | US-09-083-498-62 | Sequence 162, April | 688 | 5 | 1.2 | 254 | 51 | US-09-083-499-63 | Sequence 163, April | 689 | 5 | 1.2 | 255 | 52 | US-09-083-500-64 | Sequence 164, April | 690 | 5 | 1.2 | 256 | 53 | US-09-083-501-65 | Sequence 165, April | 691 | 5 | 1.2 | 257 | 54 | US-09-083-502-66 | Sequence 166, April | 692 | 5 | 1.2 | 258 | 55 | US-09-083-503-67 | Sequence 167, April | 693 | 5 | 1.2 | 259 | 56 | US-09-083-504-68 | Sequence 168, April | 694 | 5 | 1.2 | 260 | 57 | US-09-083-505-69 | Sequence 169, April | 695 | 5 | 1.2 | 261 | 58 | US-09-083-506-70 | Sequence 170, April | 696 | 5 | 1.2 | 262 | 59 | US-09-083-507-71 | Sequence 171, April | 697 | 5 | 1.2 | 263 | 60 | US-09-083-508-72 | Sequence 172, April | 698 | 5 | 1.2 | 264 | 61 | US-09-083-509-73 | Sequence 173, April | 699 | 5 | 1.2 | 265 | 62 | US-09-083-510-74 | Sequence 174, April | 700 | 5 | 1.2 | 266 | 63 | US-09-083-511-75 | Sequence 175, April | 701 | 5 | 1.2 | 267 | 64 | US-09-083-512-76 | Sequence 176, April | 702 | 5 | 1.2 | 268 | 65 | US-09-083-513-77 | Sequence 177, April | 703 | 5 | 1.2 | 269 | 66 | US-09-083-514-78 | Sequence 178, April | 704 | 5 | 1.2 | 270 | 67 | US-09-083-515-79 | Sequence 179, April | 705 | 5 | 1.2 | 271 | 68 | US-09-083-516-80 | Sequence 180, April | 706 | 5 | 1.2 | 272 | 69 | US-09-083-517-81 | Sequence 181, April | 707 | 5 | 1.2 | 273 | 70 | US-09-083-518-82 | Sequence 182, April | 708 | 5 | 1.2 | 274 | 71 | US-09-083-519-83 | Sequence 183, April | 709 | 5 | 1.2 | 275 | 72 | US-09-083-520-84 | Sequence 184, April | 710 | 5 | 1.2 | 276 | 73 | US-09-083-521-85 | Sequence 185, April | 711 | 5 | 1.2 | 277 | 74 | US-09-083-522-86 | Sequence 186, April | 712 | 5 | 1.2 | 278 | 75 | US-09-083-523-87 | Sequence 187, April | 713 | 5 | 1.2 | 279 | 76 | US-09-083-524-88 | Sequence 188, April | 714 | 5 | 1.2 | 280 | 77 | US-09-083-525-89 | Sequence 189, April | 715 | 5 | 1.2 | 281 | 78 | US-09-083-526-90 | Sequence 190, April | 716 | 5 | 1.2 | 282 | 79 | US-09-083-527-91 | Sequence 191, April | 717 | 5 | 1.2 | 283 | 80 | US-09-083-528-92 | Sequence 192, April | 718 | 5 | 1.2 | 284 | 81 | US-09-083-529-93 | Sequence 193, April | 719 | 5 | 1.2 | 285 | 82 | US-09-083-530-94 |
|-----|---|-----|----|---|-------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|---|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|--------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|---------------------|-----|---|-----|-----|----|------------------|

REGISTRATION NUMBER: 40,518
REFERENCE/DOCKET NUMBER: (HV)

ATTORNEY/AGENT INFORMATION:

Sequence 73, Application US/09515039
Patent No. 6214599
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/515 039
CURRENT FILING DATE: 2000-03-06
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 73
LENGTH: 345
TYPE: PRT
ORGANISM: Papain
US-09-515-039-73

Query Match 2.5% Score 10; LH 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 GCSWAFSAV 206
DB 155 GCSWAFSAV 164

RESULT 9
US-08-827-171B-7
Sequence 7, Application US/08827171B
Patent No. 6254869
GENERAL INFORMATION:
APPLICANT: CAROLIN PETERSEN
APPLICANT: JIN-XING HUANG
TITLE OF INVENTION: CYCLOPAIN VACCINES, ANTIBODIES, PEPTIDES,
TITLE OF INVENTION: PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
TITLE OF INVENTION: TREATMENT, DIAGNOSIS AND
TITLE OF INVENTION: DETECTION OF
TITLE OF INVENTION: CRYPTOSPORIDIUM PARVUM
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: PETERS, VERNY, JONES & BIK A
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 KB storage
COMPUTER: PC
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect 6.0a WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/497 171B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/014 233
FILING DATE: March 27, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hana Vermy
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: (HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-1677
TELEFAX: (415) 324-1678
INFORMATION FOR SEQ ID NO: 7;
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOTIF TYPE: protein

ORIGINAL SOURCE:
ORGANISM: catfish
US-08-827-171B-7

Query Match 2.5% Score 10; LH 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 GCSWAFSAV 206
DB 155 GCSWAFSAV 164

RESULT 10
US-09-120-365-69
Sequence 69, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/120,365
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 69
LENGTH: 443
TYPE: PRT
ORGANISM: Leishmania
US-09-120-365-69

Query Match 2.5% Score 10; LH 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 GCSWAFSAV 206
DB 147 GCSWAFSAV 156

RESULT 11
US-09-515-039-69
Sequence 69, Application US/09515039
Patent No. 6214599
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/515,039
CURRENT FILING DATE: 2000-03-06
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 69
LENGTH: 443
TYPE: PRT
ORGANISM: Leishmania
US-09-515-039-69

Query Match 2.5% Score 10; LH 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 GCSWAFSAV 206
DB 147 GCSWAFSAV 156

RESULT 12
US-09-120-365-67

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-330-121R-5

Query Match 2.28, Score 9, 18 3, Length 333
Best local Similarity 100.0%; Pred. No. 0.46;
Matches 9, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 197 GSCSWAFSA 205
DB 135 GSCSWAFSA 143

RESULT 16
US-08-883-526-3
Sequence 3, Application US/08883526
Patent No. 6033893

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Guertler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Spah, Puri
TITLE OF INVENTION: NEW HUMAN CATHEPSIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER PROGRAM: FORM
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSP for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,526
FILING DATE: Horowitz
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/PROXY NUMBER: IP-0931 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: 29715
CLONE: Genbank
US-08-883-526-3

Query Match 2.28, Score 9, 18 3, Length 333
Best local Similarity 100.0%; Pred. No. 0.46;
Matches 9, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 197 GSCSWAFSA 205
DB 135 GSCSWAFSA 143

RESULT 17
US-09-120-465-65

Sequence 65, Application US/09120465
Patent No. 6103514

GENERAL INFORMATION:
APPLICANT: Natoli, Shouji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 42290-144749
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9 333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 65
LENGTH: 333
TYPE: PRP
ORGANISM: Homo sapiens cathepsin L
US-09-120-465-65

Query Match 2.28, Score 9, 18 3, Length 333
Best local Similarity 100.0%; Pred. No. 0.46;
Matches 9, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 197 GSCSWAFSA 205
DB 135 GSCSWAFSA 143

RESULT 18
US-09-515-039-65

Sequence 65, Application US/09515039
Patent No. 6214599

GENERAL INFORMATION:
APPLICANT: Natoli, Shouji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 42290-144749
CURRENT APPLICATION NUMBER: US/09/515,039
CURRENT FILING DATE: 2000-03-06
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 65
LENGTH: 333
TYPE: PRP

ORGANISM: Homo sapiens cathepsin L
US-09-515-039-65

Query Match 2.28, Score 9, 18 3, Length 333
Best local Similarity 100.0%; Pred. No. 0.46;
Matches 9, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 197 GSCSWAFSA 205
DB 135 GSCSWAFSA 143

RESULT 19
US-08-860-255A-4

Sequence 4, Application US/08860255A
Patent No. 6274346

GENERAL INFORMATION:
APPLICANT: Abdul-Meguid, Shoulin
APPLICANT: Desjardins, Renee
APPLICANT: Jansun, Cheryl
APPLICANT: Smith, Ward
APPLICANT: Zhao, Baoquan
TITLE OF INVENTION: Method of inhibiting cathepsin K
FILE REFERENCE: 050574-X1
CURRENT APPLICATION NUMBER: US/08/860,255A
CURRENT FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: 60/008,108

1 PERSONAL INFORMATION
 2 FIRST NAME: JIM
 3 LAST NAME: WATSON
 4 DATE OF BIRTH: 01/01/1950
 5 SEX: M
 6 RACE: W
 7 HEIGHT: 5'10"
 8 WEIGHT: 170
 9 HAIR: BRN
 10 EYES: BLU
 11 SKIN: FIR
 12 TATTOO: NO
 13 SCAR: NO
 14 ALCOHOL: NO
 15 DRUGS: NO
 16 MENTAL: NO
 17 PHYSICAL: NO
 18 MEDICAL: NO
 19 ALLERGIES: NO
 20 RELIGION: NO
 21 POLITICAL: NO
 22 ETHNIC: NO
 23 ANCESTRY: NO
 24 EDUCATION: NO
 25 EMPLOYMENT: NO
 26 CREDIT: NO
 27 CREDIT REPORT: NO
 28 CREDIT SCORE: NO
 29 CREDIT HISTORY: NO
 30 CREDIT INQUIRY: NO
 31 CREDIT REVIEW: NO
 32 CREDIT UPDATE: NO
 33 CREDIT DELETE: NO
 34 CREDIT REPAIR: NO
 35 CREDIT MONITOR: NO
 36 CREDIT ALERT: NO
 37 CREDIT LOCK: NO
 38 CREDIT UNLOCK: NO
 39 CREDIT RELEASE: NO
 40 CREDIT CANCEL: NO
 41 CREDIT SUSPEND: NO
 42 CREDIT RESUME: NO
 43 CREDIT REINSTATE: NO
 44 CREDIT REEVALUATE: NO
 45 CREDIT REASSESS: NO
 46 CREDIT REANALYZE: NO
 47 CREDIT RECALCULATE: NO
 48 CREDIT RECOMPUTE: NO
 49 CREDIT RECONSTRUCT: NO
 50 CREDIT RECONSTRUCTURE: NO
 51 CREDIT RECONSTRUCTURE: NO
 52 CREDIT RECONSTRUCTURE: NO
 53 CREDIT RECONSTRUCTURE: NO
 54 CREDIT RECONSTRUCTURE: NO
 55 CREDIT RECONSTRUCTURE: NO
 56 CREDIT RECONSTRUCTURE: NO
 57 CREDIT RECONSTRUCTURE: NO
 58 CREDIT RECONSTRUCTURE: NO
 59 CREDIT RECONSTRUCTURE: NO
 60 CREDIT RECONSTRUCTURE: NO

61 PERSONAL INFORMATION
 62 FIRST NAME: JIM
 63 LAST NAME: WATSON
 64 DATE OF BIRTH: 01/01/1950
 65 SEX: M
 66 RACE: W
 67 HEIGHT: 5'10"
 68 WEIGHT: 170
 69 HAIR: BRN
 70 EYES: BLU
 71 SKIN: FIR
 72 TATTOO: NO
 73 SCAR: NO
 74 ALCOHOL: NO
 75 DRUGS: NO
 76 MENTAL: NO
 77 PHYSICAL: NO
 78 MEDICAL: NO
 79 ALLERGIES: NO
 80 RELIGION: NO
 81 POLITICAL: NO
 82 ETHNIC: NO
 83 ANCESTRY: NO
 84 EDUCATION: NO
 85 EMPLOYMENT: NO
 86 CREDIT: NO
 87 CREDIT REPORT: NO
 88 CREDIT SCORE: NO
 89 CREDIT HISTORY: NO
 90 CREDIT INQUIRY: NO
 91 CREDIT REVIEW: NO
 92 CREDIT UPDATE: NO
 93 CREDIT DELETE: NO
 94 CREDIT REPAIR: NO
 95 CREDIT MONITOR: NO
 96 CREDIT ALERT: NO
 97 CREDIT LOCK: NO
 98 CREDIT UNLOCK: NO
 99 CREDIT RELEASE: NO
 100 CREDIT CANCEL: NO

101 PERSONAL INFORMATION
 102 FIRST NAME: JIM
 103 LAST NAME: WATSON
 104 DATE OF BIRTH: 01/01/1950
 105 SEX: M
 106 RACE: W
 107 HEIGHT: 5'10"
 108 WEIGHT: 170
 109 HAIR: BRN
 110 EYES: BLU
 111 SKIN: FIR
 112 TATTOO: NO
 113 SCAR: NO
 114 ALCOHOL: NO
 115 DRUGS: NO
 116 MENTAL: NO
 117 PHYSICAL: NO
 118 MEDICAL: NO
 119 ALLERGIES: NO
 120 RELIGION: NO
 121 POLITICAL: NO
 122 ETHNIC: NO
 123 ANCESTRY: NO
 124 EDUCATION: NO
 125 EMPLOYMENT: NO
 126 CREDIT: NO
 127 CREDIT REPORT: NO
 128 CREDIT SCORE: NO
 129 CREDIT HISTORY: NO
 130 CREDIT INQUIRY: NO
 131 CREDIT REVIEW: NO
 132 CREDIT UPDATE: NO
 133 CREDIT DELETE: NO
 134 CREDIT REPAIR: NO
 135 CREDIT MONITOR: NO
 136 CREDIT ALERT: NO
 137 CREDIT LOCK: NO
 138 CREDIT UNLOCK: NO
 139 CREDIT RELEASE: NO
 140 CREDIT CANCEL: NO

141 PERSONAL INFORMATION
 142 FIRST NAME: JIM
 143 LAST NAME: WATSON
 144 DATE OF BIRTH: 01/01/1950
 145 SEX: M
 146 RACE: W
 147 HEIGHT: 5'10"
 148 WEIGHT: 170
 149 HAIR: BRN
 150 EYES: BLU
 151 SKIN: FIR
 152 TATTOO: NO
 153 SCAR: NO
 154 ALCOHOL: NO
 155 DRUGS: NO
 156 MENTAL: NO
 157 PHYSICAL: NO
 158 MEDICAL: NO
 159 ALLERGIES: NO
 160 RELIGION: NO
 161 POLITICAL: NO
 162 ETHNIC: NO
 163 ANCESTRY: NO
 164 EDUCATION: NO
 165 EMPLOYMENT: NO
 166 CREDIT: NO
 167 CREDIT REPORT: NO
 168 CREDIT SCORE: NO
 169 CREDIT HISTORY: NO
 170 CREDIT INQUIRY: NO
 171 CREDIT REVIEW: NO
 172 CREDIT UPDATE: NO
 173 CREDIT DELETE: NO
 174 CREDIT REPAIR: NO
 175 CREDIT MONITOR: NO
 176 CREDIT ALERT: NO
 177 CREDIT LOCK: NO
 178 CREDIT UNLOCK: NO
 179 CREDIT RELEASE: NO
 180 CREDIT CANCEL: NO

181 PERSONAL INFORMATION
 182 FIRST NAME: JIM
 183 LAST NAME: WATSON
 184 DATE OF BIRTH: 01/01/1950
 185 SEX: M
 186 RACE: W
 187 HEIGHT: 5'10"
 188 WEIGHT: 170
 189 HAIR: BRN
 190 EYES: BLU
 191 SKIN: FIR
 192 TATTOO: NO
 193 SCAR: NO
 194 ALCOHOL: NO
 195 DRUGS: NO
 196 MENTAL: NO
 197 PHYSICAL: NO
 198 MEDICAL: NO
 199 ALLERGIES: NO
 200 RELIGION: NO
 201 POLITICAL: NO
 202 ETHNIC: NO
 203 ANCESTRY: NO
 204 EDUCATION: NO
 205 EMPLOYMENT: NO
 206 CREDIT: NO
 207 CREDIT REPORT: NO
 208 CREDIT SCORE: NO
 209 CREDIT HISTORY: NO
 210 CREDIT INQUIRY: NO
 211 CREDIT REVIEW: NO
 212 CREDIT UPDATE: NO
 213 CREDIT DELETE: NO
 214 CREDIT REPAIR: NO
 215 CREDIT MONITOR: NO
 216 CREDIT ALERT: NO
 217 CREDIT LOCK: NO
 218 CREDIT UNLOCK: NO
 219 CREDIT RELEASE: NO
 220 CREDIT CANCEL: NO

221 PERSONAL INFORMATION
 222 FIRST NAME: JIM
 223 LAST NAME: WATSON
 224 DATE OF BIRTH: 01/01/1950
 225 SEX: M
 226 RACE: W
 227 HEIGHT: 5'10"
 228 WEIGHT: 170
 229 HAIR: BRN
 230 EYES: BLU
 231 SKIN: FIR
 232 TATTOO: NO
 233 SCAR: NO
 234 ALCOHOL: NO
 235 DRUGS: NO
 236 MENTAL: NO
 237 PHYSICAL: NO
 238 MEDICAL: NO
 239 ALLERGIES: NO
 240 RELIGION: NO
 241 POLITICAL: NO
 242 ETHNIC: NO
 243 ANCESTRY: NO
 244 EDUCATION: NO
 245 EMPLOYMENT: NO
 246 CREDIT: NO
 247 CREDIT REPORT: NO
 248 CREDIT SCORE: NO
 249 CREDIT HISTORY: NO
 250 CREDIT INQUIRY: NO
 251 CREDIT REVIEW: NO
 252 CREDIT UPDATE: NO
 253 CREDIT DELETE: NO
 254 CREDIT REPAIR: NO
 255 CREDIT MONITOR: NO
 256 CREDIT ALERT: NO
 257 CREDIT LOCK: NO
 258 CREDIT UNLOCK: NO
 259 CREDIT RELEASE: NO
 260 CREDIT CANCEL: NO

Matches 9: conservative 9: mismatches 0: indels 0: gaps 0:

QY 197 CGSCWAFSA 205
 1111111111

DB 135 CGSCWAFSA 143

RESULT 22

US-08-883-526-4
 : Sequence 4, Application US/08883526
 : Patent No. 6033893

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
 APPLICANT: Guejfer, Karl J.
 APPLICANT: Corley, Neil C.
 APPLICANT: Shah, Parvi
 TITLE OF INVENTION: NEW HUMAN CALHEPSIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq, For Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/883,526
 FILING DATE: Hefexith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PR-0331 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TEXT:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 334 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: GenBank
 US-08-883-526-4

Query Match: 2.28; Score 9; DB 4; Length 334;
 Best Local Similarity 100.0%; Ident. No. 0.46;

Matches 9: conservative 0: mismatches 0: indels 0: gaps 0:

QY 197 CGSCWAFSA 205
 1111111111

DB 135 CGSCWAFSA 143

RESULT 23

US-09-120-365-64
 : Sequence 64, Application US/09120365
 : Patent No. 6103514
 : GENERAL INFORMATION:
 APPLICANT: Natori, Shunji
 TITLE OF INVENTION: NEW PROTEASE
 : FILE REFERENCE: 32290 14474

CURRENT APPLICATION NUMBER: 32290/14474, 465
 : CURRENT FILING DATE: 1998-07-22
 : EARLIER APPLICATION NUMBER: JP 9-144 474
 : EARLIER FILING DATE: 1997-11-18
 : NUMBER OF SEQ ID NOS: 101
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 64
 : LENGTH: 334
 : TYPE: PRT
 : ORGANISM: Murine calhepsin L
 US-09-120-365-64

Query Match: 2.28; Score 9; DB 4; Length 334;
 Best Local Similarity 100.0%; Ident. No. 0.46;

Matches 9: conservative 0: mismatches 0: indels 0: gaps 0:

QY 197 CGSCWAFSA 205
 1111111111

DB 135 CGSCWAFSA 143

RESULT 24

US-09-515-039-64
 : Sequence 64, Application US/09515039
 : Patent No. 6214599

GENERAL INFORMATION:

APPLICANT: Natori, Shunji
 : TITLE OF INVENTION: NEW PROTEASE
 : FILE REFERENCE: 32290-14474
 : CURRENT FILING DATE: 1998-07-22
 : CURRENT APPLICATION NUMBER: JP 9-144 474
 : EARLIER FILING DATE: 1997-11-18
 : NUMBER OF SEQ ID NOS: 101
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 64
 : LENGTH: 334
 : TYPE: PRT
 : ORGANISM: Murine calhepsin L
 US-09-515-039-64

Query Match: 2.28; Score 9; DB 4; Length 334;
 Best Local Similarity 100.0%; Ident. No. 0.46;

Matches 9: conservative 0: mismatches 0: indels 0: gaps 0:

QY 197 CGSCWAFSA 205
 1111111111

DB 135 CGSCWAFSA 143

RESULT 25

US-09-500-651-2
 : Sequence 2, Application US/09500651
 : GENERAL INFORMATION:

APPLICANT: ASANO, MINAO
 APPLICANT: KAWAI, MITSUO
 APPLICANT: KIMURA, TOSIYUKA
 APPLICANT: KUROKI, RIKI
 TITLE OF INVENTION: NOVEL METHOD OF
 : HYDROLYSING A PROTEIN WITH THE SAME
 : NUMBER OF SEQUENCES: 8
 : CORRESPONDENCE ADDRESS:
 ADDRESSEE: P.O. BOX, HIRAKAWA, MATSUYAMA, 770-0044
 STREET: 1755 ST. JEFFERSON DAVIS HIGHWAY, SUITE 400
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC DOS/MS-DOS

Db 138 CMAFSAVA 146

RESULT 28

US-09-120-365-72
Sequence 72, Application US/09120365
Patent No. 6101514

GENERAL INFORMATION:
APPLICANT: Natoli, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US 09/120,365
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 72
LENGTH: 457
TYPE: PRI
ORGANISM: Oryzain
US-09-120-365-72

Query Match: 2.2% Score 92 DB 2 Length 457
Best Local Similarity 100.0% Prod. No. 0.61
Matches 92 Conservative 0 Mismatches 0 Indels 0 Gaps 0

UY 197 CGSCWAFSA 205
Db 150 CGSCWAFSA 158

RESULT 29

US-09-515-039-72
Sequence 72, Application US/09515039
Patent No. 6214599
GENERAL INFORMATION:
APPLICANT: Natoli, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/515,039
CURRENT FILING DATE: 2000-03-06
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 72
LENGTH: 457
TYPE: PRI
ORGANISM: Oryzain
US-09-515-039-72

Query Match: 2.2% Score 92 DB 4 Length 457
Best Local Similarity 100.0% Prod. No. 0.61
Matches 92 Conservative 0 Mismatches 0 Indels 0 Gaps 0

UY 197 CGSCWAFSA 205
Db 150 CGSCWAFSA 158

RESULT 30

US-08-119-361-15
Sequence 15, Application US/08119361
Patent No. 5523390
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
TITLE OF INVENTION: Porphyromonas gingivalis Arginine specific Protease
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5470 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS/MS 7.05
SOFTWARE: PatentIn Release #1.0, Version #1.25
CHECKED ATTACHED DATA:
ATTACHED 2 REMOTE: 95/09/11/9, 461
FILING DATE: 10-SEP-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Forbort, Donna M.
REGISTRATION NUMBER: 48878
REFERENCE/1-SECT NUMBER: 21-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TITLE: amino acid
STRANDNESS: single
TOPLOGY: linear
MOLECULE TYPE: protein
HYPOTHECAL: No
FRAGMENT TYPE: Internal
US-08-119-361-15

Query Match: 2.0% Score 81 DB 1 Length 14
Best Local Similarity 100.0% Prod. No. 0.25
Matches 8 Conservative 0 Mismatches 0 Indels 0 Gaps 0

UY 199 SCWAFSAV 206
Db 1 SCWAFSAV 8

RESULT 31
US-08-336-408A-15
Sequence 15, Application US/08336408A
Patent No. 6017542

GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine specific Protease Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5470 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.00
CURRENT APPLICATION DATA:
ALLOCATION NUMBER: 95/09/11/9, 461
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001, 411
 FILING DATE: 24 JUN 1994
 ATTORNEY/AGENT: THE SMITH M
 NAME: Robert M
 PUBLICATION NUMBER: 5,600,000
 REFERENCE/BACKL NUMBER: 21,900
 REFERENCE/INFORMATION: 21,900
 TELEPHONE: (603) 400-8000
 TELEFAX: (603) 400-8000
 INFORMATION FOR THE INVENTOR:
 INVENTOR: CHAVATHIRATHIR
 FILING DATE: 24 JUN 1994
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLGY: linear
 MOLECULE TYPE: protein
 HYDROPHOBIC: NO
 FRAGMENT TYPE: individual
 US ON 400,000,100

Query Match: 2.00% Score of 100.00 Length 14
 Best Local Similarity: 100.00% Prod. No. 0.00
 Match: M2: Conserved by 0% Mismatch: 0% Indels: 0% Gaps: 0%

09 100 SWAFSAV 200
 10 111111
 10 1 SWAFSAV M

RESULT 42
 US ON 400,000,100
 Sequence 10: Application US 04/00941
 Patent No. 42/4718
 GENERAL INFORMATION:
 APPLICANT: Francis J. M.
 APPLICANT: Robert M.
 APPLICANT: Robert M.
 APPLICANT: Robert M.
 TITLE OF INVENTION: Polysaccharide composition
 TITLE OF INVENTION: Polysaccharide composition
 NUMBER OF SEQUENCES: 10
 CORRESPONDENT ADDRESS:
 ADDRESS: Greenway, Winter and Sullivan, P.C.
 STREET: 1400 Massachusetts Blvd. and
 CITY: Boston
 STATE: Colorado
 COUNTRY: US
 ZIP: 80004
 MOLECULAR REMARKS:
 MOLECULE TYPE: Polysaccharide
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC/MS-DOS
 SOFTWARE: Polysaccharide
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Robert M.
 REFERENCE/BACKL NUMBER: 21,900
 REFERENCE/INFORMATION:
 TELEPHONE: (603) 400-8000
 TELEFAX: (603) 400-8000
 INFORMATION FOR THE INVENTOR:
 INVENTOR: CHAVATHIRATHIR

LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLGY: linear
 MOLECULE TYPE: protein
 HYDROPHOBIC: NO
 FRAGMENT TYPE: individual
 US ON 400,000,100

Query Match: 2.00% Score of 100.00 Length 14
 Best Local Similarity: 100.00% Prod. No. 0.00
 Match: M2: Conserved by 0% Mismatch: 0% Indels: 0% Gaps: 0%

09 100 SWAFSAV 200
 10 111111
 10 1 SWAFSAV M

RESULT 44
 US ON 400,000,100
 Sequence 10: Application US 04/00941

GENERAL INFORMATION:
 APPLICANT: Francis J. M.
 APPLICANT: Robert M.
 APPLICANT: Robert M.
 APPLICANT: Robert M.
 TITLE OF INVENTION: Polysaccharide composition
 TITLE OF INVENTION: Polysaccharide composition
 NUMBER OF SEQUENCES: 207
 CORRESPONDENT ADDRESS:
 ADDRESS: Greenway, Winter and Sullivan, P.C.
 STREET: 1400 Massachusetts Blvd. and
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02114
 MOLECULAR REMARKS:
 MOLECULE TYPE: Polysaccharide
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC/MS-DOS
 SOFTWARE: ASPET TEXT
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 CLASSIFICATION: 400

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 05/09/2001
 FILING DATE: 24 JUN 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Robert M.
 REFERENCE/BACKL NUMBER: 21,900
 REFERENCE/INFORMATION:
 TELEPHONE: (603) 400-8000
 TELEFAX: (603) 400-8000
 INFORMATION FOR THE INVENTOR:
 INVENTOR: CHAVATHIRATHIR
 FILING DATE: 24 JUN 1994
 TYPE: amino acids
 STRANDEDNESS: single
 MOLECULE TYPE: protein
 HYDROPHOBIC: NO
 FRAGMENT TYPE: individual
 US ON 400,000,100

Query Match: 2.00% Score of 100.00 Length 14
 Best Local Similarity: 100.00% Prod. No. 0.00
 Match: M2: Conserved by 0% Mismatch: 0% Indels: 0% Gaps: 0%

OY 197 CUSCWAFS 204
DB 19 CUSCWAFS 26

RESULT 34
US-08-482-147-74
Sequence 73, Application US/08482142
Patent No. 5920822
GENERAL INFORMATION:
APPLICANT: Garmen, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-Chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Ze'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ANTIGENS
TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE POST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,142
FILING DATE: 07-JUN 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05/08/445,307
FILING DATE: 07-June 1995
ATTORNEY/AGENT INFORMATION:
NAME: CRAIG, ANNE I.
REGISTRATION NUMBER: 32,976
REFERENCE/BOOKLET NUMBER: 017,605
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-482-147-74
Query Match 2.0%; Score 8; Db 2; Length 27;
Best Local Similarity 100.0%; Prod. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CUSCWAFS 204
DB 19 CUSCWAFS 26

RESULT 35
US-08-478-572-10
Sequence 10, Application US/08478572
Patent No. 5968526
GENERAL INFORMATION:
APPLICANT: Garmen, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-Chang

APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Ze'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ANTIGENS
TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE POST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,572
FILING DATE: 07-June-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/445,407
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CRAIG, ANNE I.
REGISTRATION NUMBER: 32,976
REFERENCE/BOOKLET NUMBER: 017,605
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-478-572-10
Query Match 2.0%; Score 8; Db 2; Length 27;
Best Local Similarity 100.0%; Prod. No. 1.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 CUSCWAFS 204
DB 19 CUSCWAFS 26

RESULT 36
US-08-478-572-74
Sequence 73, Application US/08478572
Patent No. 5968526
GENERAL INFORMATION:
APPLICANT: Garmen, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-Chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Ze'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ANTIGENS
TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE POST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA

MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-482-142-74

Query Match: 2.08; Score 8; DB 4; Length 27;
Best Local Similarity: 100.0%; Prod. No. 0.46;
Matches: 8; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0

QY 197 CGSCWAPS 204
DB 19 CGSCWAPS 26

RESULT 39
PCT-0595-04481-2
Sequence 2, Application PC/TUS9504481
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dermatitis
NUMBER OF SEQUENCES: 54
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PC/TUS95/04481
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/227,772
FILING DATE: April 14, 1994
AUTOREF/AGENT INFORMATION:
NAME: Vanstone, Darlene A.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 017.5 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
PCT-0595-04481-2

Query Match: 2.08; Score 8; DB 5; Length 27;
Best Local Similarity: 100.0%; Prod. No. 0.46;
Matches: 8; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0

QY 197 CGSCWAPS 204
DB 19 CGSCWAPS 26

RESULT 40
US-08-482-142-11
Sequence 11, Application US/08482142
GENERAL INFORMATION:
APPLICANT: Garman, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Zee'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLELGENS
TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE POST MITE)
FROM DERMATOPHAGOCYTES (HOUSE POST MITE)

NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: LINCOLN PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/209,742,142
FILING DATE: 07-JUN-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,445,407
FILING DATE: 07-June 1995
AUTOREF/AGENT INFORMATION:
NAME: CANAL, ARNE I.
REGISTRATION NUMBER: 42,976
REFERENCE/DOCKET NUMBER: 017.608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-482-142-11

Query Match: 2.08; Score 8; DB 2; Length 29;
Best Local Similarity: 100.0%; Prod. No. 0.46;
Matches: 8; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0

QY 197 CGSCWAPS 204
DB 11 CGSCWAPS 18

RESULT 41
US-08-482-142-74
Sequence 74, Application US/08482142
GENERAL INFORMATION:
APPLICANT: Garman, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Zee'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLELGENS
TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE POST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESSEE: LINCOLN PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: ASCII TEXT

QY 197 CGSWAFS 204
 Db 11 CGSWAFS 18

RESULT 44

US-08-484-296-11

Sequence 11, Application US/202421296
 Patent No. 6268491

GENERAL INFORMATION:

APPLICANT: Garman, Richard
 APPLICANT: Greenstein, Julia

APPLICANT: Kuo, Mei-Chang

APPLICANT: Rogers, Bruce

APPLICANT: Franzen, Henry

APPLICANT: Chen, Xian

APPLICANT: Evans, Sean

APPLICANT: Shaked, Zeev

TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE DUST MITE)

NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:

ADDRESS: IMMUNOGENIC PHARMACEUTICAL CORPORATION

STREET: 610 LINCOLN STREET

CITY: WALTHAM

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/202421296

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/445,307

FILING DATE: 07 June 1995

ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976

REFERENCE/DOCKET NUMBER: 017,608

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6040

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-484-296-11

Query Match 2.0%; Score 8; Db 4; Length 29;
 Best Local Similarity 100.0%; Pred. No. 9.4%;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSWAFS 204
 Db 11 CGSWAFS 18

RESULT 45

US-08-484-296-74

Sequence 74, Application US/08484296
 Patent No. 6268491

GENERAL INFORMATION:

APPLICANT: Garman, Richard
 APPLICANT: Greenstein, Julia

APPLICANT: Kuo, Mei-Chang

APPLICANT: Rogers, Bruce

APPLICANT: Franzen, Henry

APPLICANT: Chen, Xian

APPLICANT: Evans, Sean

APPLICANT: Shaked, Zeev

TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE DUST MITE)

NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:

ADDRESS: IMMUNOGENIC PHARMACEUTICAL CORPORATION

STREET: 610 LINCOLN STREET

CITY: WALTHAM

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/202421296

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/445,307

FILING DATE: 07 June 1995

ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976

REFERENCE/DOCKET NUMBER: 017,608

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6040

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-484-296-74

Query Match 2.0%; Score 8; Db 4; Length 29;
 Best Local Similarity 100.0%; Pred. No. 9.4%;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSWAFS 204
 Db 11 CGSWAFS 18

RESULT 46

US-09-598-062-4

Sequence 2, Application US/095980624
 Patent No. 6268491

GENERAL INFORMATION:

APPLICANT: Garman, Richard
 APPLICANT: Greenstein, Julia

APPLICANT: Kuo, Mei-Chang

APPLICANT: Rogers, Bruce

APPLICANT: Franzen, Henry

APPLICANT: Chen, Xian

APPLICANT: Evans, Sean

APPLICANT: Shaked, Zeev

TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

TITLE OF INVENTION: FROM DERMATOPHAGOCYTES (HOUSE DUST MITE)

NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:

ADDRESS: IMMUNOGENIC PHARMACEUTICAL CORPORATION

STREET: 610 LINCOLN STREET

CITY: WALTHAM

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/202421296

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/445,307

FILING DATE: 07 June 1995

ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976

REFERENCE/DOCKET NUMBER: 017,608

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6040

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-484-296-74

Query Match 2.0%; Score 8; Db 4; Length 29;
 Best Local Similarity 100.0%; Pred. No. 9.4%;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 CGSWAFS 204
 Db 11 CGSWAFS 18


```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc-ibm/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,570
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-21-4-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-003-570-24

```

```

Query Match          2.0%; Score 8; DB 4; Length 46;
Best Local Similarity 100.0%; Prod. No. 0.75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 CGSCWAFS 204
DB 2 CGSCWAFS 9

```

```

RESULT 50
US-08-463-262A-13
Sequence 13, Application US/08/63262A
Patent No. 5691186
GENERAL INFORMATION:
APPLICANT: TRIPP, Cynthia Ann
APPLICANT: Frank, Glenn R.
APPLICANT: Grieco, Robert B.
TITLE OF INVENTION: NOVEL FILARIN CYSTEINE PROTEINASE GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: SHERIDAN ROSS & MCINTOSH
STREET: 1700 LINCOLN STREET, SUITE 3500
CITY: DENVER
STATE: CO
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,262A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-21-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-262A-13

```

```

Query Match          2.0%; Score 8; DB 1; Length 47;

```

```

Best Local Similarity 100.0%; Prod. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 197 CGSCWAFS 204
DB 4 CGSCWAFS 10

```

```

Search completed on 10 Oct 9, 2002, 11:49:22
Job time: 30 secs

```


